

Response to reviewer #1's comments on Grise and Davis (2020)

We would like to thank the reviewer for taking time to review our manuscript and providing helpful comments. Based on the reviewer's comments, we have made a number of minor changes and clarifications to the manuscript. Detailed point-by-point responses to all comments are provided below, and the original reviewer's comments are provided in bold type.

The submission by Davis and Grise sounds at first like turn-the-crank research paper. Indeed, the analysis is a repeat of earlier work by the authors, only with some newer datasets (CMIP6 and ERA5). However, the authors do a commendable job of contrasting their results with their earlier work and under light of other recent studies. In doing so, the authors point out several outstanding questions, making this work a useful step forward. I have a few mostly-editorial comments.

We thank the reviewer for this recognition of our efforts to characterize, discuss, and interpret the CMIP6 results in light of previous findings.

Line 15-17: the sentence "First, both. . . , but this. . ." would be more clear as "First, while. . . , this. . ." to make it clear that "First" does not refer to the first clause, but to the second.

Thanks. Following the reviewer's suggestion, we have corrected this sentence to the following on lines 15–17 in the revised manuscript:

"First, while both CMIP5 and CMIP6 models contract the NH summertime Hadley circulation equatorward (particularly over the Pacific sector), this contraction is larger in CMIP6 models due to their higher average climate sensitivity."

Line 43 and 52: there's another "First" "second" list here, but it's not clearly introduced as a list, and it sounds like the list may continue afterward. "For example" and "In addition" might be better.

Thanks. We have changed "first" to "for example" (line 44) and "second" to "additionally" (line 62).

Lines 125-128: This is not the Hadley circulation boundary. If you believe the EDJ to be meaningfully related to the HC edge in the different ocean basins, you should state so, and somehow justify your belief.

Given the comments from Reviewers #1 and #2, we have eliminated the usage of the EDJ metric from the manuscript, as it does not directly correspond to the location of the Hadley cell edge (although it is highly correlated with it in terms of

interannual variability and its response to climate change; see Waugh et al. 2018). We now use the USFC metric to quantify longitudinal variability in the tropical edge. The results are very similar over the Pacific sector, which is the focus of our discussion (see new Fig. 3 and Fig. S2).

Line 145: "drastic" is the wrong word; "dramatic" is better

Thanks. We have changed “drastic” to “dramatic” on line 168 in the revised manuscript.

Line 166: Table 1 does not “support” the fact that the only significant differences occur during JJA. That fact doesn’t need supporting. But Table 1 helps explain the difference.

We have changed “is supported by” to “is consistent with” on line 189 in the revised manuscript.

A more definitive way of explaining this difference would be to normalize the shifts by the respective sensitivities (or remove the component explained by the sensitivity) and determining whether the difference remains significant once the impact of sensitivity is removed.

In the SH, where the JJA Hadley cell edge shifts are all of the same sign and of a similar order of magnitude, the reviewer’s suggestion works well. If the SH JJA Hadley cell edge shifts are divided by the global-mean surface temperature increase in each model, the difference between CMIP5 and CMIP6 models is no longer statistically significant, supporting our argument in the text and confirming the reviewer’s idea.

In the NH, the JJA Hadley cell edge shifts are of varying sign (poleward and equatorward) and of varying orders of magnitude (near-zero to almost 10°), so applying a simple normalization procedure is not straightforward to interpret (i.e., you are dividing large negative Hadley cell shifts by large positive climate sensitivities but you are dividing small negative, near-zero, and large positive Hadley cell shifts by smaller positive climate sensitivities). If we confine our analysis to only those models with equatorward Hadley cell edge shifts greater than 0.5° latitude, if the NH JJA Hadley cell edge shifts in these models are divided by the global-mean surface temperature increase, the difference between CMIP5 and CMIP6 models is no longer statistically significant. However, this result breaks down once models with near-zero and poleward Hadley cell edge shifts are included.

We thank the reviewer for the suggestion of normalizing the Hadley cell shifts by global-mean surface temperature, but because it is not straightforward to apply in the NH, we choose not to include these results in the paper.

As the reviewer also suggests, one can also use linear regression analysis to remove the variance in the Hadley cell edge shifts associated with the variance in climate sensitivity across models, but this analysis can only be used to remove the variance associated with the climate sensitivity. It does not provide any information about the contribution of the climate sensitivity to the mean Hadley cell edge shift in CMIP5 and CMIP6 models, and thus it cannot be used to assess whether the mean Hadley cell edge shifts in CMIP5 and CMIP6 models are related to the difference in mean climate sensitivity.

One can, however, compare the linear regression fits between the global-mean surface temperature response and the Hadley cell edge shifts in CMIP5 and CMIP6 models. The linear regression lines have very similar slopes in both NH JJA (approximately -1.5° latitude/Kelvin for CMIP5 and -1.25° latitude/Kelvin for CMIP6) and SH JJA (approximately -0.3° latitude/Kelvin and -0.2° latitude/Kelvin from CMIP6), suggesting that the greater mean climate sensitivity in CMIP6 models contributes to the greater dynamical sensitivity during JJA in both hemispheres. This can clearly be seen in Fig. 2a for NH JJA, as the scatter of points from both CMIP5 and CMIP6 models generally falls along the same diagonal line from the upper left toward lower right.

Line 168: Fig. 2 vs Figure 2a

We don't understand the reviewer's comment here. The first sentence introduces Fig. 2 as a whole, whereas the second sentence discusses specifics only in panel a of Fig. 2. We believe that the text is correct as written. Additionally, per ACP guidelines, the abbreviation "Fig." is used when a figure is referenced within a sentence, whereas the word "Figure" is spelled out at the beginning of a sentence.

Line 289: "with forcing" would be clearer as "with a higher sensitivity to"

We apologize that our initial wording was confusing. We are actually not discussing the relationship with climate sensitivity here, but are instead referring to the difference between the greenhouse-gas only runs and the full historical runs. We have added a parenthetical note "(compare orange, black, and red lines in Figs. 5a-b)" to clarify this to the reader on line 318 in the revised manuscript.

Line 346: "when which" should be "in which" or "during which"

We have changed "when which" to "at which" on line 374 of the revised manuscript.

Response to reviewer #2's comments on Grise and Davis (2020)

We would like to thank the reviewer for taking time to review our manuscript and providing helpful comments. Based on the reviewer's comments, we have made a number of minor changes and clarifications to the manuscript. Detailed point-by-point responses to all comments are provided below, and the original reviewer's comments are provided in bold type.

The manuscript documents and contrasts trends in the Hadley cell edge in current generation CMIP6 models, the previous generation CMIP5 models and multiple reanalysis products. The results are new, novel and will be of broad interest to the community. Put simply the manuscript is superb and I very little reservation in recommending this for publication with only a few minor changes. The manuscript is skillfully written and clearly communicated. I particularly liked the perfectly simple and concise title, an excellent and well-motivated introduction, and plots of a consistently high quality and common style. The manuscript also has a nice balance of both the NH and SH perspectives. I would be happy to review the paper a second time should it be needed.

We thank the reviewer for her very positive assessment of our manuscript.

In the introduction the focus is very much Hadley cell edge rather than tropical edge. So I was a bit surprised then to see the EDJ metric used in the analysis. I think a more general description of Hadley cell edge vs tropical edge metrics is then needed in the introduction. Then bring this through to your second last paragraph in the introduction that describes the aims of the paper.

As discussed below, we have eliminated the usage of the EDJ metric in the analysis. As such, we believe that the aims of the paper in the Introduction are correct as stated.

My interpretation for the use of the EDJ is that your taking advantage of the correlations between the HC edge and the EDJ in order to get longitudinal variability. Can you explain why you used this approach over say a local HC method (eg Schwendike et al. (2014) or your own study Staten et al. (2019))? Is this to avoid adding more metrics (ie consistent with TropD and the idea that fewer/better methods are the goal) or is it that you have concerns about the local methods accuracy/applicability?

Given the comments from Reviewers #1 and #2, we have eliminated the usage of the EDJ metric from the manuscript, as it does not directly correspond to the location of the Hadley cell edge (although it is highly correlated with it). We now use the USFC metric to quantify longitudinal variability in the tropical edge. We appreciate the reviewer's recognition of our previous

work that attempts to define the Hadley cell at individual longitudes through local overturning circulations. However, this concept is still relatively new and remains an area of active research, so we feel that using the USFC metric at individual longitudes will be more straightforward for readers to interpret.

To address the reviewer's concern, we have added the following text into the methods section on lines 140-145 of the revised manuscript:

"We also make brief use of the USFC metric to examine longitudinal asymmetries in the circulation response, as the PSI500 metric can only strictly be defined in the zonal mean. Some recent studies have attempted to generalize the zonal-mean Hadley cell edge (as defined by the PSI500 metric) to individual longitudes by isolating regional meridional overturning cells (Schwendike et al., 2014; Staten et al., 2019). However, interpreting these regional overturning circulations is challenging and remains an area of active research, and thus we do not examine these local overturning cells here."

Introduction: In addition to the papers already cited, there are a number of excellent papers on Hadley cell expansion from Chris Lucas and Hanh Nguyen at the Bureau of Meteorology which could also be cited here. I am thinking papers such as Nguyen et al. (2012, 2015, 2018) Lucas et al. (2012, 2013); Lucas and Nguyen (2015). I am not suggesting you cite them all, but one or two would be good.

We have added citations to Lucas et al. (2012), Lucas et al. (2014), and Nguyen et al. (2015) in the introduction (lines 33, 37, 52). Many of their other papers focus on the tropopause height metric of tropical width, which does not co-vary interannually with the Hadley cell edge (e.g., Waugh et al. 2018), and/or regional aspects of tropical widening, which are not the focus of the text in the introduction.

The introduction would benefit from a brief discussion that the Hadley cell latitude (middle atmosphere mass stream function zero crossing latitude) is thought to be the most reliable measure of tropical expansion (this is touched on briefly in L45 but I think this could be expanded a little). I think there should also be mention to other tropical edge metrics (STJ, EDJ, OLR, P-E etc) which are used within the literature already cited, and include the relationships with the jets (strength and position) in terms of tropical edge metrics and relationships with the HC position Ceppi and Hartmann (2012); Menzel et al. (2019); Maher et al. (2019).

We have added the following text into the introduction to more fully describe the metrics used to define the edges of the tropics (lines 45–58 of the revised manuscript):

"Traditionally, the edge of the Hadley circulation has been defined using the poleward boundary of the zonal-mean meridional mass streamfunction in the mid-troposphere, but departures from mass conservation in reanalyses (particularly in older

generation reanalyses) can lead to large spurious trends in the location of the Hadley cell edge defined using the mass streamfunction (Davis and Davis, 2018). Consequently, many studies have sought to estimate trends in the location of the Hadley cell edge using other metrics, including the transition from zonal-mean surface easterlies to zonal-mean surface westerlies (Grise et al., 2018, hereafter G18; Grise et al., 2019, hereafter G19), the subtropical sea level pressure maximum (Choi et al., 2014), the latitude of the subtropical jet (Maher et al., 2020), the altitude break in tropopause height in the subtropics (Seidel and Randel, 2007; Lucas et al., 2012), thresholds in outgoing longwave radiation (Hu and Fu, 2007; Mantsis et al., 2017), and total column ozone (Hudson et al., 2006). Some of the largest trends in recent decades arise from the metrics derived from tropopause height and outgoing longwave radiation, but it appears that these metrics are measuring changes unrelated to the poleward expansion of the Hadley circulation. While all of the metrics listed above co-locate climatologically with the poleward boundary of the mass streamfunction, only the surface wind and sea level pressure metrics co-vary interannually with the streamfunction boundary (Davis and Birner, 2017; Davis et al., 2018; Solomon et al., 2016; Waugh et al., 2018), at least in reanalyses and models.”

We do not feel that it is necessary to provide a detailed discussion of the relationships among the strengths and positions of the subtropical and eddy-driven jets, particularly because we no longer use the EDJ metric in the manuscript. The focus of the manuscript is on the Hadley cell edge, not on the jets.

L21: I found ‘indicate’ awkward in this sentence. I suggest rearranging to the following or similar: ‘The poleward expansion of the Hadley circulation is one of the most robust aspects of the atmospheric general circulations response to a warming climate in global circulation models.’

Following the reviewer’s suggestion, we have changed the first sentence of the introduction (lines 21-22 of the revised manuscript) to the following:

“The poleward expansion of the Hadley circulation is one of the most robust aspects of the atmospheric general circulation’s response to a warming climate in global climate models.”

The acronyms CMIP, SH and NH are already written in full in the abstract so I do not think you need to define them again in the introduction.

According to the ACP manuscript preparation guidelines, abbreviations “need to be defined in the abstract and then again at the first instance in the rest of the text.” So, per the guidelines, we define the acronyms in both the abstract and main text.

L79-90: I think this would be easier to read directly from Table S1/S2 rather than list in the paragraph. I would then

move the tables from the supplementary into the manuscript. Instead of ‘x’ you could add the time window of the data, add a column for the reference for each model, add a column for indicating if CMIP5 or 6 (then you only need 1 table) and include the horizontal resolution of the model. This would probably then be a landscape whole page table which is common for CMIP papers. These are simply suggestions, proceed as you wish.

Following the reviewer’s suggestion, we have removed Tables S1 and S2 and placed the salient information into a new table (Table 1) in the main text. We have also added the horizontal resolution of each model into the table, as the reviewer requested. However, we do not include the references for each of the 44 individual models used in this study, as we found it difficult to ascertain the appropriate references for some of the models. Referring to the citation requirements stated on the CMIP website (<https://pcmdi.llnl.gov/CMIP6/TermsOfUse/TermsOfUse6-1.html>), we follow their recommendation to cite the relevant articles published in the CMIP6 special issue of GMD. No mention is made of citing the papers from the individual modeling centers.

To keep Table 1 relatively concise, we have chosen to retain the information about the time periods of the runs in the first paragraph of section 2.1. However, we have rewritten this paragraph to include an enumerated list of the different model runs (lines 91-97), which will hopefully make this paragraph easier to read.

L107-112: I would also use a table for the reanalysis data sets.

The details of the reanalysis data sets are now listed in Table 2.

It would be mentioning in this section the method used to test significance (it is stated in each of the plots already).

We had added the following paragraph at the end of section 2b to address how we calculate statistical significance (lines 146-151 of the revised manuscript):

“We evaluate whether the multi-model means of CMIP5 and CMIP6 models are statistically different from one another using a two-tailed Student’s t-test. When comparing values from CMIP5 and CMIP6 models, we use large asterisks in the figures to denote where the multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence level. For the significance testing, we treat each model as an independent sample. However, because many climate models are closely related to one another (e.g., Knutti et al., 2013), the actual value of significance is likely to be much lower.”

Fig 1 (and S1): I think this plot would look better and take up less space if it were 1 row and 2 column (ie side by side).

Following the reviewer's suggestion, we have reformatted Fig. 1, Fig. S1, and Fig. S5 in a 1 row-2 column format.

I found the asterisk (and in later plots the circles) hard to interpret (are they in S1 too?). I think open and filled circles for the ensemble mean might communicate this clearer or have a lighter and darker versions of black/red. Likewise for Fig 4

We have added the following text into section 2b (lines 147-149 of the revised manuscript) to clarify the meaning of the asterisks in the figures:

“When comparing values from CMIP5 and CMIP6 models, we use large asterisks in the figures to denote where the multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence level.”

We appreciate the reviewer's suggestion, but this would not be straightforward to apply in Figures 6–7, where two different significant tests are applied (i.e., testing whether the mean trend in both CMIP5 and CMIP6 models is statistically different from zero and testing whether the multi-model means of CMIP5 and CMIP6 models are statistically different). We prefer to use consistent symbols and formatting across all figures, so we prefer to retain the use of the asterisks to denote where the multi-model means of CMIP5 and CMIP6 models are statistically different.

The large dots in Figs. 6–7 follow the convention of Fig. 2 from G19 to test whether the mean trend is statistically different from zero, so we prefer to follow the same format to allow comparison with our previous study.

Was the goal of Fig 2-3 on focusing on the NH only to draw out the differences seen in Fig 1 for JJA in NH?

Yes. We now clarify in the text (lines 191-192 of the revised manuscript) why we focus on the NH JJA response in Figs. 2–3:

“we further examine the largest difference between CMIP5 and CMIP6 models identified in Fig. 1: the response of the NH JJA Hadley cell edge to 4xCO₂ forcing”

Fig 5: The legends are a little small and repeated. Suggest making larger and only having one (perhaps at the top or bottom of the panel). Might also help to add 'ensemble mean' for the purple line in the top panels of each sub-plot.

Following the reviewer's suggestions, we have eliminated the repeated legends in Figure 5, enlarged the text of the remaining legend in panel d, and added a legend entry denoting the “multi-reanalysis mean” in the top panel of Figure 5d.

L238: only 3 out of the models ‘greatly exceed’ historical and AMIP runs. Suggest mentioning the reanalysis models by name or putting in the clause ‘some of the reanalysis greatly exceed’.

Per the reviewer’s request, we have corrected this sentence to be more precise (lines 264-266 of the revised manuscript):

“For the PSI500 metric (Fig. 4, left column), trends from the ERA-Interim, MERRA-2, and JRA-55 reanalyses in the NH and from the ERA-Interim reanalysis in the SH are substantially larger than the trends from the models’ control runs and greatly exceed the trends from the historical and AMIP runs of most models (see also G18, G19).”

Fig 6-7: I can’t really tell one model from another. I am not sure this level of detail is helpful as I found these plots a little overwhelming. The ensemble mean bars are also easily lost in the scatter. This is my personal opinion and the authors can change or not change these figures as they wish.

In Figs. 6 and 7 (also Fig. S3), we have lightened the symbols for the individual models, so that the ensemble mean bars are more easily visible. We agree with the reviewer that there is a lot of information in these figures, but we are updating a similar figure from a prior study (Fig. 2 of Grise et al. 2019). We show the symbols for individual models for two reasons. First, for the common models that performed both CMIP5 and CMIP6 single forcing runs, it allows the reader to see whether the circulation response to a particular forcing notably changed between the CMIP5 and CMIP6 versions of that model. Second, because different models performed single forcing runs in CMIP5 and CMIP6, it allows the reader to assess which models may be contributing to the differences in the CMIP5 and CMIP6 ensemble mean responses to a particular forcing.

What is happening in Fig c-d bottom panels for the reanalysis between 1990-2000 – is this the PDO?

The equatorward anomalies in the SH Hadley cell edge in the early 1990s followed by the poleward anomalies in the SH Hadley cell edge in the late 1990s are consistent with the change in phase of the PDO from positive to negative. The AMIP runs of the models also capture this feature to a lesser extent, showing a pause in the poleward SH Hadley cell edge trend in the early 1990s followed by an acceleration of the poleward SH Hadley cell edge trend in the late 1990s. However, this feature is much larger in the reanalyses, likely because concurrent internal atmospheric variability also contributes to large decadal variability in the Hadley cell edge. Unlike coupled atmosphere-ocean variability, the timing of internal atmospheric variability is not necessarily the same in the AMIP runs of the models as in observations, allowing for notable deviations in the decadal variability of the reanalyses and models seen in Fig. 5.

While this feature is interesting, the focus of Section 4 is on the long-term trends, not on the decadal variability in the reanalysis

time series. For this reason, we choose not to discuss the decadal variability in the observed Hadley cell edge in the paper.

Suggest starting this section with ‘The’ so that the 5 and 21 are separated.

We have changed the title of section 5 to be “Projected Hadley cell expansion over the 21st century” (see line 365 of the revised manuscript).

The first page would benefit from adding a title of the paper and stating it is the supplementary material (minimum or adding title page (if you wish)).

According to the ACP manuscript preparation guidelines, “supplements will receive a title page added during the publication process including title (“Supplement of”), authors, and the correspondence email. Therefore, please avoid providing this information in the supplement.” So, per the guidelines, we have not included a title page to the supplement.

Response to reviewer #3's comments on Grise and Davis (2020)

We would like to thank the reviewer for taking time to review our manuscript and to provide helpful comments. Based on the reviewer's comments, we have made a number of minor changes and clarifications to the manuscript. Detailed point-by-point responses to all comments are provided below, and original reviewers' comments are provided in bold type.

General comments: This is a very thorough and thoughtful study on the expansion of the Hadley cell across CMIP5 models/CMIP6 models/reanalyses. It clearly lays out the key similarities and differences between CMIP5 and CMIP6, and sets the differences between the models and reanalysis in a useful context. I have one main comment, and if it is addressed, the manuscript would be suitable for publication.

We thank the reviewer for their overall positive assessment of our manuscript and efforts to usefully compare CMIP5 models, CMIP6 models, and reanalyses.

My comment is that the role of the pattern of SST warming for the NH JJA contraction should be discussed. For CMIP5 models, there are significant differences in Hadley cell edge response between the amipFuture and amip4K experiments for the NH JJA season. So it would be helpful if results from the amip4K experiment were shown for comparison. Furthermore, Zhou et al. (2019) argue that an ITCZ shift related to the pattern of enhanced equatorial SST warming drives the subtropical circulation contraction during NH JJA. This is different from the present manuscript, in which it is argued that a general SST warming (pattern not required) forces the equatorial contraction during summer.

As requested by the reviewer, we have added results from the amip4K experiments to Figure 2c (also to Figure S2b). The reviewer is correct that the sign of the NH Hadley cell edge response changes when the uniform 4K SST warming is used instead of the patterned 4K SST warming. To clarify this, we have added the following text on lines 215-217 of the revised manuscript:

"However, as pointed out by Zhou et al. (2019), the exact pattern of SST warming is critical for capturing the equatorward contraction of the NH JJA Hadley cell edge seen in the abrupt 4xCO₂ runs. A uniform 4K SST warming would instead result in a poleward expansion of the NH JJA Hadley circulation (Fig. 2c)."

We note that our original text largely followed from Shaw and Voigt (2015)'s results. They concluded that both the amip4K and amipFuture runs contribute to an equatorward contraction of the circulation in the Pacific basin, with the amip4K changes being slightly weaker. A subsequent careful examination of their paper shows that they exclude the western Pacific from their analysis (see red box in their Fig. 1). When averaging over the entire Pacific basin (and thus also in the zonal mean) as we do

here, a poleward expansion of the summertime circulation in the western Pacific basin is sufficient to overwhelm the equatorward contraction of the circulation in the eastern Pacific basin in the amip4K runs, but not in the amipFuture runs. This accounts for the apparent contradiction in the results of Shaw and Voigt (2015), who argue that the amip4K and amipFuture runs contribute to the same sign of the circulation response, and Zhou et al. (2019), who argue that the amip4K and amipFuture runs contribute to different signs of the circulation response.

1) How do you define the PSI500 Hadley cell edge for NH JJA if PSI never becomes positive in the tropics, and hence a zero-crossing does not exist? In my experience this occurs for some years in certain models.

The reviewer is correct that the NH JJA Hadley cell edge is undefined during some years. We have added the following text to the methods section (lines 132-136 of the revised manuscript) to clarify this to the reader:

“We note that the NH summertime Hadley circulation is very weak, making it challenging to define the PSI500 metric during some years. We only consider the PSI500 metric from years in which there is a clear crossing of the 500-hPa streamfunction field from positive to negative in the NH subtropics. We consider the PSI500 metric to be undefined if no zero crossing in the streamfunction field occurs or if multiple zero crossings from positive to negative occur within a 20° latitude band (‘Lat_Uncertainty = 20’ in TropD).”

2) How exactly is the "response" defined for the abrupt4xCO2 experiment (Fig. 1)? I.e. over what period of the abrupt4xCO2 experiment are you averaging?

The reviewer must have missed this definition, which is stated in the second sentence of section 3 (lines 155-157): “difference in the Hadley cell edge latitude between its mean position during the last 50 years (years 101-150) of the abrupt 4xCO₂ run and its mean position in the pre-industrial control run.”

We have now also added this definition to the caption of Fig. 1 (lines 699-701) to make sure readers are aware of how the quantity plotted in Fig. 1 is defined.

References: Zhou, W., S.-P. Xie and D. Yang (2019): Enhanced equatorial warming causes deertropical contraction and subtropical monsoon shift. Nature Climate Change. 9. 834- 839.

Thanks. We have added a citation to this paper (lines 663-664).

List of changes to the manuscript

Abstract

- Following reviewer #1's suggestion, we have reworded the sentence on lines 15–17.

Section 1

- Following reviewer #2's suggestion, we have reworded the first sentence of the introduction (lines 21-22 of the revised manuscript).
- We have added references to Lucas et al. (2012), Lucas et al. (2014), and Nguyen et al. (2015) in response to reviewer #2's comment.
- Following reviewer #1's suggestion, we have changed “first” to “for example” (line 44) and “second” to “additionally” (line 62).
- In response to reviewer #2's comment, we have added text into the introduction to more fully describe the metrics used to define the edges of the tropics (lines 45–58 of the revised manuscript).

Section 2

- We have rewritten the first paragraph of section 2.1 in response to reviewer #2's comment.
- We have added a description of the amip4K runs, which we now use in response to reviewer #3's suggestion.
- We have removed the details of the reanalysis data sets from the text and placed them in Table 2 (per reviewer #2's suggestion).
- In response to reviewer #3's comment, we have added text to the methods section (lines 132-136 of the revised manuscript) to address how we deal with the Northern Hemisphere summer Hadley cell edge being poorly defined during some years.
- We have eliminated use of the eddy-driven jet metric (per the comments of Reviewers #1 and #2), and replaced it with the USFC metric to identify longitudinal asymmetries in the circulation trends.
- In response to Reviewer #2's comment, we have added text on lines 140-145 addressing the regional meridional overturning cells or local Hadley cell perspective used by some previous studies, explaining why we do not use it in this study.
- In response to Reviewer #2's comment, we had added a paragraph at the end of section 2b to address how we calculate statistical significance (lines 146-151 of the revised manuscript).

Section 3

- We have changed “drastic” to “dramatic” on line 168 in the revised manuscript, per reviewer #1's suggestion.
- We added a citation to a new paper by Zelinka et al. (2020) documenting the higher climate sensitivity in CMIP6 models.
- We have changed “is supported by” to “is consistent with” on line 189 in the revised manuscript, per reviewer #1's suggestion.
- In response to reviewer #2's comment, we now clarify in the text (lines 191-192 of the revised manuscript) why we focus on the NH JJA response in Figs. 2–3.
- As requested by reviewer #3, we have added results from the amip4K experiments to Figure 2c (also to Figure S2b) and describe the results on lines 215-217 of the revised manuscript, including a citation to the Zhou et al. (2019) paper referenced by reviewer #3.
- We have eliminated use of the eddy-driven jet metric (per the comments of Reviewers #1 and #2), and replaced it with the USFC metric to identify longitudinal asymmetries in the circulation trends. Results in Fig. 3 and S2 and

the discussion in section 3 have been updated accordingly.

Section 4

- Per reviewer #2's request, we have corrected the sentence on lines 264-266 to be more precise.
- In response to reviewer #1's confusion, we have added a parenthetical note "(compare orange, black, and red lines in Figs. 5a-b)" to clarify on line 318 in the revised manuscript.

Section 5

- In response to reviewer #2's comment, we have changed the title of section 5 to be "Projected Hadley cell expansion over the 21st century" (see line 365 of the revised manuscript).
- We have changed "when which" to "at which" on line 374 of the revised manuscript, per reviewer #1's suggestion.

Figures and Tables

- Following reviewer #2's suggestion, we have removed Tables S1 and S2 and placed the salient information into a new table (Table 1) in the main text. We have also added the horizontal resolution of each model into the table, as the reviewer requested.
- Following reviewer #2's suggestion, we have added a new table (Table 2) listing the reanalysis data sets used in this study and removed the information from the text in section 2.1.
- Following reviewer #2's suggestion, we have reformatted Fig. 1, Fig. S1, and Fig. S5 in a 1 row-2 column format.
- Per reviewer #3's comment, we have added the definition of the response to the caption of Fig. 1 (lines 699-701) to make sure readers are aware of how the quantity plotted in Fig. 1 is defined.
- As requested by reviewer #3, we have added results from the amip4K experiments to Figure 2c (also to Figure S2b).
- We have eliminated use of the eddy-driven jet metric (per the comments of Reviewers #1 and #2), and replaced it with the USFC metric to identify longitudinal asymmetries in the circulation trends. Results in Fig. 3 and S2 have been updated accordingly.
- Following reviewer #2's suggestions, we have eliminated the repeated legends in Figure 5, enlarged the text of the remaining legend in panel d, and added a legend entry denoting the "multi-reanalysis mean" in the top panel of Figure 5d.
- In response to reviewer #2's comment, in Figs. 6 and 7 (also Fig. S3), we have lightened the symbols for the individual models, so that the ensemble mean bars are more easily visible.

Hadley cell expansion in CMIP6 models

Kevin M. Grise¹, Sean M. Davis²

¹Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22904, USA

²NOAA Earth System Research Laboratory Chemical Sciences Division, Boulder, CO 80305, USA

5 Correspondence to: Kevin M. Grise (kmg3r@virginia.edu)

Abstract. In response to increasing greenhouse gases, the subtropical edges of Earth's Hadley circulation shift poleward in global climate models. Recent studies have found that reanalysis trends in the Hadley cell edge over the past 30–40 years are within the range of trends simulated by Coupled Model Intercomparison Project Phase 5 (CMIP5) models, and have documented seasonal and hemispheric asymmetries in these trends. In this study, we evaluate whether these conclusions hold for the newest generation of models (CMIP6). Overall, we find similar characteristics of Hadley cell expansion in CMIP5 and CMIP6 models. In both CMIP5 and CMIP6 models, the poleward shift of the Hadley cell edge in response to increasing greenhouse gases is 2–3 times larger in the Southern Hemisphere (SH), except during September–November. The trends from CMIP5 and CMIP6 models agree well with reanalyses, although prescribing observed coupled atmosphere-ocean variability allows the models to better capture reanalysis trends in the Northern Hemisphere (NH). We find two notable differences between CMIP5 and CMIP6 models. First, while both CMIP5 and CMIP6 models contract the NH summertime Hadley circulation equatorward (particularly over the Pacific sector), this contraction is larger in CMIP6 models due to their higher average climate sensitivity. Second, in recent decades, the poleward shift of the NH annual-mean Hadley cell edge is slightly larger in CMIP6 models. Increasing greenhouse gases drive similar trends in CMIP5 and CMIP6 models, so the larger recent NH trends in CMIP6 models point to the role of other forcings, such as aerosols.

Deleted: but

20 1 Introduction

The poleward expansion of the Hadley circulation is one of the most robust aspects of the atmospheric general circulation's response to a warming climate in global climate models. This response is seen in models of varying complexity, ranging from idealized aquaplanet simulations (Frierson et al., 2007; Levine and Schneider, 2011; Tandon et al., 2013) to comprehensive general circulation model experiments (Hu et al., 2013; Lu et al., 2007; Tao et al., 2016), such as those from phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP). The poleward expansion of the Hadley circulation is anticipated to have a number of regional climate impacts in the subtropics, potentially shifting dry regions (Feng and Fu, 2013; Scheff and Frierson, 2012; Schmidt and Grise, 2017), altering zones of ocean upwelling (Cook and Vizy, 2018; Rykaczewski et al., 2015), and modifying hurricane tracks (Kossin et al., 2014; Sharmila and Walsh, 2018; Studholme and Gulev, 2018).

Deleted: Global climate models indicate that

Deleted: is the poleward expansion of the Hadley circulation.

A decade ago, a number of studies began estimating rates of Hadley cell expansion using various observational data sets (Fu et al., 2006; Hu and Fu, 2007; Seidel and Randel, 2007; Seidel et al., 2008). These rates varied widely by study, ranging from 0.2° to 3° latitude per decade over the period from 1979 until the mid-2000s (Birner et al., 2014; Davis and

35 Rosenlof, 2012; [Lucas et al., 2014](#)). The largest observed trends were an order of magnitude larger than those projected by climate models over the same period (Hu et al., 2013; Johanson and Fu, 2009), calling into question whether the observed trends were biased high and/or whether the models were deficient in simulating circulation trends. Additionally, studies disagreed on the cause of the observed trends. Some studies identified an important role for anthropogenic forcing, including increasing greenhouse gases (Hu et al., 2013; [Nguyen et al., 2015](#); Tao et al., 2016), stratospheric ozone depletion (Kang et al., 2011; McLandress et al., 2011; Min and Son, 2013; Polvani et al., 2011; Son et al., 2010), and changes in anthropogenic aerosols (Allen et al., 2012; Allen and Ajoku, 2016; Kovilakam and Mahajan, 2015). However, other studies concluded that the observed trends strongly reflected natural climate variability (Allen and Kovilakam, 2017; Amaya et al., 2018; Mantsis et al., 2017).

Deleted:).

45 Recent efforts by the US CLIVAR Working Group on the Changing Width of the Tropical Belt and the International Space Science Institute (ISSI) Tropical Width Diagnostics Intercomparison Project have addressed many of these discrepancies in the previous literature. [For example](#), the large observed rates of expansion documented by some earlier studies have been attributed to methodological issues. [Traditionally, the edge of the Hadley circulation has been defined using the poleward boundary of the zonal-mean meridional mass streamfunction in the mid-troposphere, but departures from mass conservation in reanalyses \(particularly in older generation reanalyses\) can lead to large spurious trends in the location of the Hadley cell edge defined using the mass streamfunction \(Davis and Davis, 2018\). Consequently, many studies have sought to estimate trends in the location of the Hadley cell edge using other metrics, including the transition from zonal-mean surface easterlies to zonal-mean surface westerlies \(Grise et al., 2018, hereafter G18; Grise et al., 2019, hereafter G19\), the subtropical sea level pressure maximum \(Choi et al., 2014\), the latitude of the subtropical jet \(Maher et al., 2020\), the altitude break in tropopause height in the subtropics \(Seidel and Randel, 2007; Lucas et al., 2012\), thresholds in outgoing longwave radiation \(Hu and Fu, 2007; Mantsis et al., 2017\), and total column ozone \(Hudson et al., 2006\). Some of the largest trends in recent decades arise from the metrics derived from tropopause height and outgoing longwave radiation, but it appears that these metrics are measuring changes unrelated to the poleward expansion of the Hadley circulation. While all of the metrics listed above co-locate climatologically with the poleward boundary of the mass streamfunction, only the surface wind and sea level pressure metrics co-vary interannually with the streamfunction boundary \(Davis and Birner, 2017; Davis et al., 2018; Solomon et al., 2016; Waugh et al., 2018\), at least in reanalyses and models.](#) Accounting for these issues, estimates of the recent expansion of the Hadley circulation have been narrowed to be $\leq 0.5^\circ$ latitude per decade and within the range of trends indicated by global climate models over the historical period (G18; Staten et al., 2018).

Deleted: First

Deleted: Some of the largest rates of expansion arise from measurements of tropopause height and outgoing longwave radiation, which do not co-vary interannually with the Hadley cell edge (Davis and Birner, 2017; Solomon et al., 2016; Waugh et al., 2018), at least in reanalyses and models. Using ozone measurements to define the location of the Hadley cell edge has similarly been found to be problematic (Davis et al., 2018). Furthermore, departures from mass conservation in reanalyses, particularly in older generation reanalyses, can lead to large spurious trends in the location of the Hadley cell edge (Davis and Davis, 2018).

65 [Additionally](#), in terms of the attribution of the recent trends, [G19](#) concluded that the recent poleward expansion of the Southern Hemisphere (SH) Hadley cell edge was driven in part by anthropogenic forcing (increasing greenhouse gases and stratospheric ozone depletion) and in part by natural variability, whereas the recent poleward expansion of the Northern Hemisphere (NH) Hadley cell edge was predominantly driven by natural variability. While the observed rates of expansion are approximately comparable in the two hemispheres, models indicate that anthropogenic forcing alone should drive 3–4 times larger expansion in the SH (cf. Fig. 2 of G19). Over the historical period, stratospheric ozone depletion plays a key role

Deleted: Grise et al., 2018, hereafter

Deleted: Second

Deleted: Grise et al. (2019, hereafter G19)

85 in this hemispheric asymmetry, especially during austral summer (DJF). However, even in models forced only by increasing greenhouse gases, the poleward shift of the SH Hadley cell edge is substantially larger than that in the NH (Davis et al., 2016; Grise and Polvani, 2016; Watt-Meyer et al., 2019); only during the SON season are expansion rates comparable between the two hemispheres. G19 concluded that the role of aerosols in the observed Hadley cell expansion appears to be small based on CMIP5 models, but remains very uncertain due to the diverse treatment of aerosols in models.

90 Most of the conclusions discussed above were formulated using CMIP5 model output, and as CMIP represents an “ensemble of opportunity,” it is quite possible that some of the relationships established from CMIP5 models may have been unique to that model generation. The goal of this study is to re-evaluate key conclusions about Hadley cell expansion in a new generation of global climate models (CMIP6) and to assess their robustness across model generation. CMIP6 includes output from updated versions of CMIP5 models (many of which have different treatments of clouds and aerosols, among other factors), as well as new models that did not participate in CMIP5. Overall, we find that the characteristics of Hadley cell expansion are very similar in CMIP5 and CMIP6 models, but we find several notable exceptions, which we detail below.

95 The paper is organized as follows. Section 2 details the data and methods. Section 3 examines the response of the Hadley cell edge to an idealized 4xCO₂ forcing in CMIP6 models and compares the results to CMIP5 models. Section 4 then examines the trends from the historical runs of CMIP6 models, and contrasts them with reanalyses and CMIP5 models. Section 5 briefly compares the 21st century trends in CMIP5 and CMIP6 models. Section 6 provides a summary and concluding thoughts.

2 Data and Methods

2.1 Data

105 The primary data used in this study are output from the 24 CMIP5 (Taylor et al., 2012) and 20 CMIP6 (Eyring et al., 2016) models, listed in Table 1. These models were selected because they had data available from all of the following runs at the time of the writing of this manuscript:

- 1) pre-industrial control; fully coupled runs simulating 200+ years of unforced variability.
- 2) historical; fully coupled runs forced with observed radiative forcings over the period 1850–2005 for CMIP5 and 1850–2014 for CMIP6.
- 110 3) AMIP; atmosphere-only runs forced with observed radiative forcings, sea surface temperatures, and sea ice concentrations over the period 1979–2008 for CMIP5 and 1979–2014 for CMIP6.
- 4) abrupt 4xCO₂; fully coupled runs in which atmospheric CO₂ concentrations are abruptly quadrupled from pre-industrial levels and held fixed for 150 years.

_____ Additionally, to examine a high emissions scenario for the 21st century, we use the Representative Concentration

- Deleted: model
- Deleted:). For these
- Deleted: , we use
- Deleted: :
- Formatted: Font: Italic
- Deleted: (
- Deleted:),
- Formatted: Font: Italic
- Deleted: (
- Deleted:),
- Formatted: Font: Italic
- Deleted: (
- Deleted:), and
- Formatted: Font: Italic
- Deleted: (
- Deleted:).
- Formatted: Indent: First line: 0"

Pathway (RCP) 8.5 runs (2006–2100) for CMIP5 models and the Shared Socioeconomic Pathway (SSP) 5-8.5 runs (2015–2100) for CMIP6 models. All 24 CMIP5 models have data available for the RCP 8.5 scenario, but only 14 of the 20 CMIP6 models have data available for the SSP 5-8.5 scenario (see CMIP6 models marked with # symbol in Table 1).

130 For a subset of the models in Table 1, we use three additional runs, which are useful in the attribution of Hadley cell expansion. Following Grise and Polvani (2014), we use the amip4xCO₂ and amipFuture (called “amip-future4K” for CMIP6) runs to partition the circulation response to increased atmospheric CO₂ into components associated with the direct radiative forcing of CO₂ (amip4xCO₂ – AMIP) and sea surface temperature (SST) warming (amipFuture – AMIP). The amip4xCO₂ runs are atmosphere-only runs with the same SSTs and sea ice as the AMIP runs, but with quadrupled atmospheric CO₂ concentrations; the amipFuture runs add a patterned SST anomaly (normalized to a global-mean value of 4K) to the AMIP SSTs, but retain the same CO₂ and sea ice concentrations as the AMIP runs (Webb et al., 2017). To determine whether the results are sensitive to the patterned SST anomaly used in the amipFuture runs, we also examine the amip4K (called “amip-p4K” for CMIP6) runs, which add a uniform SST anomaly of 4K to the AMIP SSTs, but retain the same CO₂ and sea ice concentrations as the AMIP runs (Webb et al., 2017). 10 CMIP5 models and 7 CMIP6 models have output available for the amip4xCO₂, amipFuture, and amip4K runs (see bolded models in Table 1).

140 Over the historical period (1850–2005 for CMIP5, 1850–2014 for CMIP6), single forcing runs are also examined from available models (see Table S1 for CMIP5 and Table S2 for CMIP6). These runs are identical to the historical runs, except that they only prescribe one forcing over the historical period: well-mixed greenhouse gases, natural (solar and volcanic), anthropogenic aerosols, and ozone. Note that, in CMIP5 models, the ozone-only runs include changes in both stratospheric and tropospheric ozone concentrations, whereas the ozone-only runs in CMIP6 models are only forced by changes in stratospheric ozone concentrations. Furthermore, some CMIP5 models included ozone changes in their greenhouse gas only runs (Gillett et al., 2016), and following G19, we exclude those models here to more clearly separate the influences of stratospheric ozone depletion and increasing greenhouse gases on the circulation response.

150 To compare the historical circulation trends in models with observations, we make use of the five modern reanalysis data sets, listed in Table 2. Because the CFSR reanalysis ends in 2010, we extend it through 2014 using CFSv2. We do not examine the NCEP-NCAR or NCEP-DOE reanalyses here, as they contain substantial departures from mass conservation over the historical period (Davis and Davis, 2018).

2.2 Methods

155 To locate the edges of the Hadley circulation, we make use of two metrics: PSI500 and USFC. PSI500 is defined as the subtropical latitude where the zonal-mean meridional mass streamfunction at 500 hPa switches sign from thermally direct (Hadley circulation) to thermally indirect (Ferrel circulation). USFC is defined as the subtropical latitude where the zonal-mean zonal wind at the surface switches sign from tropical easterlies to midlatitude westerlies. The metrics are calculated using the Tropical-width Diagnostics code package (TropD; Adam et al., 2018). Before calculating these metrics, the wind fields are zonally and time averaged (i.e., annual-mean, zonal-mean or seasonal-mean, zonal-mean wind fields are used). We

Deleted: For CMIP5 models, 24 available models have data from all of these runs (see Table S1). For CMIP6 models, 20 available models have data from the control, historical, AMIP, and abrupt 4xCO₂ runs at the time of the writing of this manuscript, with 14 of those models also having output from the SSP 5-8.5 runs (see Table S2). Unless otherwise noted, the first ensemble member is used for each model. ...

Deleted: these

Deleted: several

Deleted: 10 CMIP5 models and 7 CMIP6 models have output available for the amip4xCO₂ and amipFuture runs (see Tables S1 and S2). ...

Deleted: S3

Deleted: S4

Deleted: : 1) ERA-Interim (1979–2014; Dee et al., 2011),

Deleted:) NASA MERRA-2 (1980–2014; Gelaro et al., 2017), 3) JRA-55 (1979–2014; Kobayashi et al., 2015), 4) NCEP CFSR (1979–2010; Saha et al., 2010), and 5) ERA-5 (1979–2014; Hersbach et al., 2019).

Deleted: (Saha et al., 2014).

180 note that the NH summertime Hadley circulation is very weak, making it challenging to define the PSI500 metric during some
years. We only consider the PSI500 metric from years in which there is a clear crossing of the 500-hPa streamfunction field
from positive to negative in the NH subtropics. We consider the PSI500 metric to be undefined if no zero crossing in the
streamfunction field occurs or if multiple zero crossings from positive to negative occur within a 20° latitude band
(‘Lat_Uncertainty = 20’ in TropD).

185 In this paper, we focus on results for the PSI500 metric, as it is the most widely used metric of Hadley cell width in
the previous literature. Key results for the USFC metric are shown in the supplementary material. However, when comparing
the Hadley cell expansion in models with observations, we show results from both metrics, because of potential biases in the
PSI500 metric in reanalyses (Davis and Davis, 2018; G19). We also make brief use of the USFC metric to examine longitudinal
asymmetries in the circulation response, as the PSI500 metric can only strictly be defined in the zonal mean. Some recent
190 studies have attempted to generalize the zonal-mean Hadley cell edge (as defined by the PSI500 metric) to individual
longitudes by isolating regional meridional overturning cells (Schwendike et al., 2014; Staten et al., 2019). However,
interpreting these regional overturning circulations is challenging and remains an area of active research, and thus we do not
examine these local overturning cells here.

195 We evaluate whether the multi-model means of CMIP5 and CMIP6 models are statistically different from one another
using a two-tailed Student’s t-test. When comparing values from CMIP5 and CMIP6 models, we use large asterisks in the
figures to denote where the multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence
level. For the significance testing, we treat each model as an independent sample. However, because many climate models are
closely related to one another (e.g., Knutti et al., 2013), the actual value of significance is likely to be much lower.

Deleted: To examine longitudinal asymmetries in the circulation response, we also make brief use of a metric for the eddy-driven jet position over the North Pacific and North Atlantic Ocean basins. The eddy-driven jet position is defined as the latitude of the maximum in the 850-hPa zonal wind field, which is zonally averaged over the longitude range of each ocean basin (135°E-125°W for North Pacific, 60°W-0°E for North Atlantic). Following G18, the eddy-driven jet latitude is calculated using TropD, using the ‘max’ option in the function TropD_Metric_EDJ.

3 Dynamical sensitivity of CMIP6 models

200 Before examining Hadley cell expansion over the historical period, we first compare and contrast the dynamical
sensitivity of CMIP5 and CMIP6 models. Following Grise and Polvani (2016, hereafter GP16), we define dynamical
sensitivity as the response of the circulation to 4xCO₂ forcing, which is calculated here as the difference in the Hadley cell
edge latitude between its mean position during the last 50 years (years 101-150) of the abrupt 4xCO₂ run and its mean position
in the pre-industrial control run. Examining the dynamical sensitivity is important, as it directly allows us to compare
205 generations of models to a common forcing. The abrupt 4xCO₂ experiment is chosen for this purpose, as it is a standard
experiment planned to be included in all future phases of CMIP (Eyring et al., 2016). In contrast, the forcings used in the
historical and future scenario runs of CMIP models change across model generations, making it difficult to verify whether
differences between model generations are because of model improvements or changes in forcings.

Figure 1 shows the response of the NH and SH Hadley cell edge latitudes (as measured by the PSI500 metric) to
210 4xCO₂ forcing. Qualitatively similar results for the USFC metric are shown in the supplementary material (Fig. S1). In the
SH, both CMIP5 and CMIP6 models show ~2° of Hadley cell expansion in response to 4xCO₂ forcing. The SH expansion has

little variation across the seasonal cycle, with slightly larger poleward shifts of the Hadley cell edge in MAM and SON (see also GP16). On average, the poleward expansion seen in CMIP6 models is only slightly larger than that in CMIP5 models, with the difference between CMIP5 and CMIP6 models only being statistically significant in JJA.

In the NH, the response of the Hadley cell edge to 4xCO₂ has a more **dramatic** seasonal variation. In the annual mean, the multi-model mean Hadley cell expansion is $\sim 0.75^\circ$ latitude, roughly 40% of the multi-model mean response in the SH. The smaller poleward shift of the NH Hadley cell edge in the annual mean reflects a compensation between a large poleward shift of the NH Hadley cell edge in SON (and to a lesser extent in DJF) and a large equatorward shift of the NH Hadley cell edge in JJA. This seasonality is consistent with previously published results based on CMIP5 models (GP16; Watt-Meyer et al., 2019). The differences between CMIP5 and CMIP6 models are small in all seasons except JJA. However, in JJA, the equatorward contraction of the circulation is notably larger in CMIP6 models. As a result, three CMIP6 models (CESM2, CESM2-WACCM, and SAM0-UNICON) actually contract the annual-mean NH Hadley cell edge equatorward, a result not seen in CMIP5 models (at least as measured by the PSI500 metric).

The differences between CMIP5 and CMIP6 models in Fig. 1 may be because the CMIP6 models, on average, have a higher climate sensitivity (Forster et al., 2019; Zelinka et al., 2020). To check this, in Table 3, we show correlations between the annual-mean global-mean surface temperature response to 4xCO₂ forcing and the Hadley cell edge response across the inter-model spread of both CMIP5 and CMIP6 models. The results support the conclusions of GP16 based upon CMIP5 models. In the SH, the magnitude of the poleward shift in the Hadley cell edge is strongly correlated with the global-mean surface temperature response throughout the year, with the largest and most significant correlations in MAM and JJA (cf. Fig. 4 of GP16). In other words, models that warm more in response to 4xCO₂ forcing tend to shift the SH Hadley cell edge further poleward. In contrast, in the NH, the magnitude of the shift in the Hadley cell edge is very poorly correlated with the global-mean surface temperature response in the annual mean. This largely reflects a compensation between a significant positive correlation in DJF and a significant negative correlation in JJA. That is, models that warm more in response to 4xCO₂ forcing tend to shift the NH Hadley cell edge further poleward in DJF but also further equatorward in JJA. The fact that the only significant differences between CMIP5 and CMIP6 models in Fig. 1 occur in the JJA season in both hemispheres is **consistent** with Table 3, as JJA is the season with the largest magnitude correlation between the dynamical sensitivity and the global-mean surface temperature response in both hemispheres.

In Fig. 2, we further examine the **largest difference between CMIP5 and CMIP6 models identified in Fig. 1: the response of the NH JJA Hadley cell edge to 4xCO₂ forcing**. Figure 2a shows the scatter plot between the responses of the global-mean surface temperature and the NH JJA Hadley cell edge latitude to 4xCO₂ forcing. As documented in Table 3, the strong anti-correlation between the NH JJA Hadley cell edge shift and the global-mean surface temperature response is clearly visible. Because CMIP6 models have on average 1 K greater warming in response to 4xCO₂ forcing (6.1 K for CMIP6, compared to 5.1 K for CMIP5), the NH JJA Hadley cell edge shifts significantly further equatorward (-4° latitude for CMIP6, compared to 1.5° latitude for CMIP5).

The timeseries of the response of the NH JJA Hadley cell edge latitude to an abrupt quadrupling of atmospheric CO₂

Deleted: drastic

Deleted:).

Deleted: 1

Deleted: supported by

Deleted: 1

Deleted: 1

yields further insight into the processes involved (Fig. 2b). Initially, in both CMIP5 and CMIP6 models, the Hadley cell edge shifts slightly poleward in the first decade after CO₂ quadrupling, but then retreats equatorward for the remainder of the 150-year run. Consistent with Figs. 1 and 2a, the equatorward retreat of the NH JJA Hadley cell edge is substantially larger in CMIP6 models.

265 Following Grise and Polvani (2014) and Shaw and Voigt (2015), we can examine the roles of the direct radiative effects of CO₂ and SST warming in this circulation response (see methods in Sect. 2a). In response to a quadrupling of atmospheric CO₂ concentrations (but no change in SSTs), both CMIP5 and CMIP6 models show a ~0.6° latitude poleward expansion of the NH JJA Hadley circulation (Fig. 2c), consistent with the immediate circulation response in Fig. 2b after abrupt CO₂ quadrupling. In contrast, both CMIP5 and CMIP6 models show a ~1.0° latitude equatorward contraction of the
270 NH JJA Hadley circulation in response to a patterned 4K SST warming (with no change in atmospheric CO₂ concentrations). NH summer is the season when circulation changes driven by the direct radiative effects of CO₂ most clearly oppose those driven by SST warming (Grise and Polvani, 2014). As argued by Shaw and Voigt (2015), the direct radiative effects of CO₂ enhance land-sea temperature contrast and act to shift the circulation poleward, whereas the SST warming reduces land-sea temperature contrast and acts to shift the circulation equatorward. Because the SST warming is larger in CMIP6 models on
275 average (due to their higher climate sensitivity), the SST-driven component of the circulation response would be expected to be larger in CMIP6 models, resulting in a larger net equatorward contraction of the NH Hadley circulation during JJA than in CMIP5 models. However, as pointed out by Zhou et al. (2019), the exact pattern of SST warming is critical for capturing the equatorward contraction of the NH JJA Hadley cell edge seen in the abrupt 4xCO₂ runs. A uniform 4K SST warming would instead result in a poleward expansion of the NH JJA Hadley circulation (Fig. 2c).

280 One may question the meaningfulness of looking at the NH summertime Hadley circulation, which is generally very weak (Dima and Wallace, 2003) and largely reflects regional overturning circulations in the Indian Ocean/West Pacific sector (Hoskins et al., 2019). So, to aid in the interpretation of the results in Figs. 1–2, we also examine the regional structure of the NH circulation response during JJA. Figure 3a shows the multi-model mean surface zonal wind response to 4xCO₂ forcing for the JJA season for CMIP6 models. From this figure, it is clear that the equatorward contraction of the NH summertime
285 circulation arises largely from the Pacific sector, consistent with findings from CMIP5 models (Shaw and Voigt, 2015; GP16).

There is little net shift in the subtropical surface wind field over the Atlantic sector during JJA, (see also Fig. 3c).

The latitude of the transition between tropical surface easterlies and midlatitude surface westerlies over the North Pacific shifts poleward in most seasons but shifts equatorward in summer (Fig. 3b), similar to the zonal-mean Hadley circulation (Fig. 1). In CMIP6 models, the winter and fall circulation shifts further poleward over the Pacific sector than in
290 the CMIP5 models, but the summer circulation shifts further equatorward. As a result, there is little difference in annual-mean circulation shifts between CMIP5 and CMIP6 models over either the North Pacific or North Atlantic sectors. As noted above for the zonal-mean Hadley circulation (Fig. 2), the equatorward contraction of the Pacific circulation during JJA results from the competing effects of the direct radiative effects of CO₂ and SST warming on the circulation (see Fig. S2). The equatorward contraction of the Pacific circulation is larger on average in CMIP6 models (Fig. 3b), as the effect of the warming SSTs

Deleted: 850 hPa

Deleted: In contrast,

Deleted: North

Deleted: circulation expands poleward

Deleted: . In fact, the eddy-driven jet over the North Atlantic shifts poleward in all seasons but DJF (

Deleted:), whereas the eddy-driven jet

Deleted: winter

Deleted:)

Deleted: wintertime

Deleted: significantly

Deleted: summertime

Deleted: significantly

Deleted: jet

Deleted: notably

310 overpowers any poleward expansion driven by the direct radiative effects of CO₂.

In summary, in this section, we compared and contrasted the responses of the NH and SH Hadley cell edges to abrupt 4xCO₂ forcing. The magnitudes and seasonality of the Hadley cell expansion in CMIP6 models are very similar to those in CMIP5 models (Fig. 1). The most notable differences occur in the JJA season, particularly in the NH where CMIP6 models show a substantially larger equatorward contraction of the circulation than CMIP5 models. During this season, the response of the NH Hadley cell edge to 4xCO₂ forcing is significantly anti-correlated with the global-mean surface temperature response (Table 3; Fig. 2a), and because the average climate sensitivity of CMIP6 models is larger, the circulation contracts further equatorward in CMIP6 models. This equatorward contraction of the NH Hadley cell during summer largely reflects an equatorward shift of the circulation over the Pacific sector (Fig. 3), where there is a competition between the direct radiative effects of CO₂ (which act to expand the circulation poleward) and SST warming (which acts to contract the circulation equatorward). Because the CO₂ forcing is the same but the SST warming is larger in CMIP6 models, the net equatorward contraction of the NH summertime circulation is notably larger in CMIP6 models.

Deleted: 1

4 Hadley cell expansion over the historical period

Having compared the models' Hadley cell edge response to a common forcing, we now use this knowledge to compare the models' behavior over the historical period. Figure 4 shows the trends in the annual-mean Hadley cell edge latitude (as measured by both the PSI500 and USFC metrics) over the period 1979–2008 from five reanalyses, CMIP5 models, and CMIP6 models. We examine this 30-year period as it represents the common period covered by the AMIP runs of both CMIP5 and CMIP6 models. Because CMIP5 models' historical runs end in 2005, we have extended these runs with three years of the RCP 8.5 runs until 2008. Qualitatively similar results are found if slightly different end dates are used instead of 2008. For reference, in Fig. 5, we plot the reanalysis and multi-model mean timeseries from which the trends in Fig. 4 are calculated.

Figure 4 shows that the observed trends for the USFC metric (as estimated by reanalyses) are relatively modest ($\leq 0.2^\circ$ latitude per decade in each hemisphere) and within the bounds of the 30-year trends from the control runs of the models (see also G18, G19). In the NH, the reanalysis trends lie at the upper range of trends from the models' historical runs and fall near the multi-model mean trend from the models' AMIP runs, suggesting an important role for SST variability in driving the recent poleward shift in the NH Hadley cell edge (Allen et al., 2014; Allen and Kovilakam, 2017; G19). In the SH, the reanalysis trends compare well with the multi-model mean trends from the historical runs of CMIP5 and CMIP6 models and the multi-model mean trend from the AMIP runs of CMIP5 models. The multi-model mean trend from the AMIP runs of CMIP6 models compares well with the trend from the ERA-5 reanalysis but exceeds the trends from the other reanalyses.

For the PSI500 metric (Fig. 4, left column), trends from the ERA-Interim, MERRA-2, and JRA-55 reanalyses in the NH and from the ERA-Interim reanalysis in the SH are substantially larger than the trends from the models' control runs and greatly exceed the trends from the historical and AMIP runs of most models (see also G18, G19). As discussed by G19, the PSI500 metric is subject to considerable uncertainty in reanalyses (see spread in reanalysis time series in Fig. 5) because of

Deleted: the reanalysis

Deleted: in

Deleted: NH (and also

Deleted:

Deleted:)

inconsistencies in assimilated satellite radiances across reanalyses (Fujiwara et al., 2017) and departures from mass conservation (Davis and Davis, 2018). By contrast, at least some of the surface pressure and marine surface wind observations are shared among reanalysis centers (Fujiwara et al., 2017), resulting in stronger agreement among the reanalysis time series for the USFC metric (Fig. 5, right column).

Over the 1979–2008 period, the trends from the historical and AMIP runs of CMIP5 and CMIP6 models are very similar, with two key exceptions. First, as noted above, for the USFC metric, the trends in the SH Hadley cell edge are significantly larger in the AMIP runs of CMIP6 models than in the AMIP runs of CMIP5 models (Fig. 4d), but this result is metric dependent and does not hold for the PSI500 metric (Fig. 4c). Second, for both the PSI500 and USFC metrics, the trends in the NH Hadley cell edge are significantly larger in the historical runs of CMIP6 models than in the historical runs of CMIP5 models (Figs. 4a-b). This can also clearly be seen in the time series in Fig. 5 and is not unique to the 1979–2008 period highlighted in Fig. 4. The discrepancy between the historical trends in CMIP5 and CMIP6 models in Fig. 4 is unexpected, as increased CO₂ results in very similar trends in the NH annual-mean Hadley cell edge in CMIP5 and CMIP6 models (Fig. 1). Indeed, CMIP6 models forced only with increasing greenhouse gases over the historical period (Fig. 5, orange lines) compare very favorably with the historical runs of CMIP5 models (Fig. 5, solid black lines). This evidence suggests that other forcings (solar/volcanic, aerosol, ozone) could be contributing to the larger NH circulation trends in recent decades in the historical runs of CMIP6 models.

To address the role of different forcings in contributing to trends in the models' historical runs, we examine trends in the Hadley cell edge latitude from all available ensemble members of the historical single forcing runs of CMIP5 and CMIP6 models, updating the results of G19 to include CMIP6 models (see their Fig. 2). Results for the NH Hadley cell edge latitude are shown in Fig. 6, and results for the SH Hadley cell edge latitude are shown in Fig. 7. Recall that these single forcing runs are only available from a small subset of the models (8 CMIP5 models and 9 CMIP6 models, as listed in Tables S1 and S2). Following G19, results are shown for two time periods, 1950–2005 and 1979–2005, where 1950 is the start year of the single forcing runs in some CMIP5 models and 2005 is the end year of the single forcing runs in CMIP5 models.

In the NH, CMIP5 and CMIP6 models agree that increasing greenhouse gases were the dominant forcing contributing to a poleward shift of the annual-mean Hadley cell edge over the second half of the 20th century (Fig. 6). However, the poleward trends in the Hadley cell edge latitude in the NH associated with increasing greenhouse gases are ~2–3 times smaller than those in the SH, consistent with the results from the abrupt 4xCO₂ runs shown in Fig. 1. The roles of the remaining forcings (solar/volcanic, aerosol, ozone) are smaller and are of inconsistent sign between CMIP5 and CMIP6 models. Natural (solar/volcanic) forcing contributes to a poleward shift of the NH Hadley cell edge over the 1979–2005 period in CMIP5 models (Allen et al., 2014), but an equatorward shift of the NH Hadley cell edge over the same period in CMIP6 models. Anthropogenic aerosol forcing contributes to a statistically significant equatorward shift of the NH Hadley cell edge over the 1950–2005 period in CMIP5 models (Allen and Ajoku, 2016), but this influence has weakened in CMIP6 models (particularly for the USFC metric). Finally, the ozone single forcing run is associated with a poleward shift of the NH Hadley cell edge in CMIP5 models (Allen et al., 2014), but not in CMIP6 models. Here, a large difference between CMIP5 and CMIP6 models

Deleted: S3

Deleted: S4

385 is expected, as the ozone single forcing runs are driven by both tropospheric and stratospheric ozone forcing in CMIP5 models
but only by stratospheric ozone forcing in CMIP6 models (which is well known to have a much larger effect on the circulation
in the SH).

390 Unfortunately, for this subset of models with single forcing runs, the difference in the historical trends in the NH
Hadley cell edge latitude between CMIP5 and CMIP6 models (Fig. 6) is smaller than for the entire ensemble of models shown
in Fig. 4. Consequently, it is difficult to use these runs to fully understand the discrepancies in the models' historical runs
shown in Figs. 4–5. For the USFC metric, the historical trends from the 9 CMIP6 models with single forcing runs are larger
than those from the 8 CMIP5 models with single forcing runs (Figs. 6c-6d), consistent with Fig. 4b. Over the 1950–2005
period, the trends in the historical runs of CMIP5 models reflect a compensation between a poleward shift of the Hadley cell
edge due to greenhouse gas forcing and an equatorward shift of the Hadley cell edge due to anthropogenic aerosol forcing
(Fig. 6c). In CMIP6 models, the aerosol influence on the circulation is weaker, allowing the greenhouse gas forcing to
395 dominate. A similar but weaker pattern in the trends is seen over the 1979–2005 period for the USFC metric (Fig. 6d), but not
for the PSI500 metric (Fig. 6b). Therefore, while Fig. 6 provides some limited evidence that aerosol forcing may play a role
in the discrepancy in the NH historical circulation trends between CMIP5 and CMIP6 models (Figs. 4–5), it is difficult to
generalize these conclusions based on a small subset of models to the entire multi-model ensemble. What is clear is that the
larger historical trends in CMIP6 models over the last several decades appear inconsistent with forcing by increasing
400 greenhouse gases alone, (compare orange, black, and red lines in Figs. 5a-b).

Figure 7 shows the trends in the SH Hadley cell edge from the historical single forcing runs of CMIP5 and CMIP6
models for the PSI500 metric for both the annual mean and the DJF season. Results for the USFC metric are shown in Fig.
S3. The results in Fig. 7 largely support the results from Fig. 2 of G19 based on CMIP5 models alone. Over the second half
of the 20th century, the models indicate that increasing greenhouse gases and stratospheric ozone depletion (particularly during
405 DJF) were the dominant forcings contributing to a poleward shift of the SH Hadley cell edge. There is also some suggestion
that anthropogenic aerosols contributed to a slight equatorward contraction of the SH Hadley cell edge, particularly over the
1950–2005 period (see also Choi et al., 2019). The one notable difference in the SH historical trends between CMIP5 and
CMIP6 models is that the circulation trends associated with the ozone forcing appear to be significantly weaker in CMIP6
models. However, only a small number of models conducted the historical ozone forcing runs, and unfortunately none of the
410 same modeling centers conducted the runs for both CMIP5 and CMIP6. Therefore, inter-model differences in the circulation
response to ozone forcing likely play a role in the discrepancy between CMIP5 and CMIP6 models seen in Fig. 7, particularly
because the magnitude of the austral spring polar lower stratospheric cooling in response to stratospheric ozone depletion is
similar in CMIP5 and CMIP6 models (not shown). The inclusion of tropospheric ozone forcing in the CMIP5 single forcing
runs may also be a factor.

415 Finally, we explore the seasonality of the recent trends in the NH and SH Hadley cell edge latitudes. Time series of
the reanalysis and multi-model mean Hadley cell edge latitudes for all four seasons are shown in Fig. 8. For reference, we
also plot the 1979–2008 trends from individual reanalyses and models in Fig. S4. Given the confounding issues with the

Deleted: .

PSI500 metric in reanalyses discussed above, we focus on the USFC metric in these figures.

420 In the NH, the reanalysis time series show near-zero to slightly equatorward trends in the Hadley cell edge during MAM and JJA and fall close to the multi-model mean of the CMIP historical runs during these seasons (see also G18). However, during DJF and SON, the reanalysis time series show sizeable (-0.3° – 0.4° latitude per decade) poleward trends in the Hadley cell edge. During these seasons, the magnitude of the reanalysis trends is larger than the trends from the historical and AMIP runs of most models (Fig. S4; see also G18). In DJF, the AMIP runs of CMIP5 and CMIP6 models approximate 425 the reanalysis trends better than the historical runs (Fig. 8), suggesting the importance of recent SST variability in driving the observed NH circulation trends during this season. In SON, the multi-model mean trends from CMIP6 models' historical and AMIP runs are larger than those from CMIP5 models and are in better agreement with the reanalysis trends (Fig. 8, compare red and black lines). Hence, the larger trends in the NH Hadley cell edge in CMIP6 models noted above in the annual mean most clearly manifest themselves during SON (compare Fig. 5b with Fig. 8).

430 In the SH, the reanalysis time series show consistent poleward trends in the Hadley cell edge (-0.2° – 0.3° latitude per decade in all seasons but JJA), falling close to the multi-model mean trends from the CMIP5 and CMIP6 historical runs during all seasons (see also G18). During all seasons but DJF, the time series of the SH Hadley cell edge from reanalyses also closely parallels the time series from the CMIP6 runs forced only by increasing greenhouse gases (compare orange and thick blue lines in the right column of Fig. 8). However, during DJF, the historical greenhouse gas only runs substantially underestimate 435 the trends in reanalyses, pointing to the importance of stratospheric ozone depletion in driving SH circulation trends during this season (Figs. 7c-d), as documented by numerous previous studies (Garfinkel et al., 2015; McLandress et al., 2011; Min and Son, 2013; Polvani et al., 2011; Son et al., 2010; Waugh et al., 2015).

In summary, in this section, we examined the trends in the latitudes of the NH and SH Hadley cell edges over the late 20th century and early 21st century in CMIP5 and CMIP6 models and compared them to trends from five reanalyses. Our 440 conclusions largely support the conclusions of recent studies documenting Hadley cell expansion in CMIP5 models (e.g., Allen and Kovilakam, 2017; G18; G19). However, we find that the historical trends in the annual-mean NH Hadley cell edge latitude are significantly larger over the 1979-2008 period in CMIP6 models (Fig. 4). One might be tempted to attribute the larger trends in CMIP6 models to their higher average climate sensitivity, but as shown in Sect. 3, the larger historical circulation trends in CMIP6 models are actually inconsistent with greenhouse gas forcing, which drives comparable magnitude shifts in 445 the NH annual-mean Hadley cell edge in CMIP5 and CMIP6 models (Figs. 1 and 6). We instead conclude that some other forcing (possibly aerosol forcing, see Fig. 6) must be contributing to the larger historical circulation trends in CMIP6 models.

5 Projected Hadley cell expansion over the 21st century.

Deleted: trends

Finally, we briefly compare the 21st century trends from the RCP 8.5 runs of CMIP5 models with those from the SSP 5-8.5 runs of CMIP6 models. Figure 9 shows the time series of the annual-mean NH and SH Hadley cell edge latitudes over 450 the period 1920–2100 based on the PSI500 metric. The time series show the multi-model mean of the historical runs extended

through the 21st century with the RCP 8.5 runs for CMIP5 models and the SSP 5-8.5 runs for CMIP6 models. The multi-model mean 20th and 21st century time series for CMIP5 and CMIP6 models are virtually identical. For reference, we provide a scatter plot of the 2015–2100 trends from individual models (as well as the trends by season) in Fig. S5. Given that the RCP 8.5 and SSP 5-8.5 runs are dominated by greenhouse gas forcing, the results in Fig. S5 are very similar to those shown for the 4xCO₂ forcing in Fig. 1, but with slightly weaker magnitude.

Following Hawkins and Sutton (2012) and G19, we define a “timescale of emergence” as the time at which the multi-model mean forced circulation response surpasses a given threshold of natural variability (as defined from the models’ control runs). For the SH, both the CMIP5 and CMIP6 multi-model mean Hadley cell edge latitudes surpass the one standard deviation threshold of variability in the models’ control runs (Fig. 9, gray shading) around the year 2000 (Fig. 9b), suggesting that the circulation response to anthropogenic forcing may have already emerged from natural variability (at least by this measure). This early emergence arises principally from the DJF season (G19; Solomon and Polvani, 2016; Thomas et al., 2015), due in large part to the added influence of stratospheric ozone depletion on the circulation during this season. In this high emissions scenario, the SH annual-mean Hadley cell edge would surpass the two standard deviation threshold of variability in the models’ control runs (Fig. 9, gray dashed lines) around the year 2045. This timescale is slightly faster than the timescale of emergence (2060) derived from the Community Earth System Model (CESM) Large Ensemble (G19; Quan et al., 2018).

In the NH, as noted by G19, the circulation response would take much longer to emerge from natural variability. In this high emissions scenario, the NH annual-mean Hadley cell edge would surpass the one standard deviation threshold of variability in the models’ control runs between 2060–2070 and would not surpass the two standard deviation threshold of variability in the 21st century (Fig. 9a). Again, this timescale is faster than that noted for the CESM Large Ensemble by G19, who did not find the poleward shift of the NH Hadley cell edge to be large enough to emerge from natural variability in the 21st century in that model. Regardless, the NH circulation response will take much longer to emerge from natural variability than the SH circulation response. This is for two reasons: 1) the larger magnitude response of the Hadley cell edge to increasing greenhouse gases in the SH (Fig. 1) and 2) the slightly larger range of natural variability in the Hadley cell edge latitude in the NH (compare gray shading in Fig. 9a and 9b). Note that, during the SON season, the poleward shift of the NH Hadley cell edge may emerge from natural variability as early as 2040 (not shown), due to the larger NH circulation response to greenhouse gas forcing during that season (Fig. 1).

6 Summary and conclusions

In response to increasing greenhouse gases, global climate models show a robust poleward expansion of the Hadley circulation (Davis et al., 2016; GP16; Watt-Meyer et al., 2019), and numerous lines of observational evidence suggest that the Hadley circulation has already expanded over the last 30–40 years (Birner et al., 2014; Davis and Rosenlof, 2012; Seidel et al., 2008; Staten et al., 2018). Within the past 5 years, studies have used output from CMIP5 global climate models to better understand the causes of the observed expansion (Allen et al., 2014; Allen and Kovilakam, 2017; G19) and to predict its

Deleted: when

485 possible evolution over the 21st century (Hu et al., 2013; Tao et al., 2016). In this paper, we assess whether these conclusions are robust across model generations by examining output from CMIP6 models.

We find strong agreement in the trends in the latitudes of the NH and SH Hadley cell edges from CMIP5 and CMIP6 models in response to abrupt 4xCO₂ (Fig. 1), historical (Fig. 4), and 21st century (Fig. 9) forcings. Specifically, we find a number of features to be robust across model generation:

- 490
- Models that warm more in response to CO₂ forcing (i.e., models with a higher climate sensitivity) generally shift the SH Hadley cell edge further poleward during all seasons, shift the NH Hadley cell edge further poleward during DJF, but contract the NH Hadley cell edge further equatorward during JJA (Table 3; GP16). The equatorward contraction of the NH circulation during summer arises from the Pacific sector (Fig. 3; Grise and Polvani, 2014; Shaw and Voigt, 2015).
- 495
- In response to CO₂ forcing, models shift the annual-mean Hadley cell edge 2–3 times further poleward in the SH than in the NH (Fig. 1; GP16; Watt-Meyer et al., 2019). Only during the SON season is the Hadley circulation expansion comparable in the two hemispheres. This implies that, with continued increases in greenhouse gases, the circulation response will emerge from natural variability in the 21st century much sooner in the SH than in the NH (Fig. 9; G19).
- 500
- Over the last 30–40 years, the magnitude of the Hadley cell expansion indicated by reanalyses using the USFC metric is within the range of trends simulated by CMIP models' historical and AMIP runs (Fig. 4; G19). Large discrepancies between reanalysis and model trends primarily result from examining trends in the PSI500 metric, which has known biases in reanalyses (Davis and Davis, 2018; G19).
- 505
- Observed coupled atmosphere-ocean variability has likely played an important role in recent trends, particularly in the NH (Fig. 4; Allen and Kovilakam, 2017). Increasing greenhouse gases and stratospheric ozone depletion have likely played an important role in recent trends in the SH (Fig. 7).

There are, however, several notable differences in CMIP6 models. First, the equatorward contraction of the NH summertime circulation is stronger in CMIP6 models, apparently as a result of their higher average climate sensitivity (Fig. 2). Second, over the last 30–40 years, the annual average trends in the NH Hadley cell edge in CMIP6 models' historical runs are slightly larger than those in CMIP5 models' historical runs. This discrepancy is not associated with differences in climate sensitivity, as trends in greenhouse-gas only runs over this time period agree well between CMIP5 and CMIP6 models (Figs. 5–6). The biggest discrepancies in historical circulation trends between CMIP5 and CMIP6 models appear to arise from other forcings (solar/volcanic, anthropogenic aerosol, ozone), which contribute to substantial variance in circulation trends across model generations (Figs. 6–7).

515 Overall, there is good agreement on the characteristics of Hadley circulation expansion in CMIP5 and CMIP6 models, yet several outstanding issues remain that require further understanding. First, the consistency of the hemispheric and seasonal asymmetries of the circulation trends across model generation attests to their robustness, emphasizing a greater need to better

Deleted: 1

understand the physical mechanisms responsible for these asymmetries (see discussion in Watt-Meyer et al., 2019). Second, a better understanding is needed of the roles of non-greenhouse gas forcings on historical circulation trends and why these trends diverge significantly across model generation. Finally, we have focused almost entirely on zonal-mean circulation trends in this paper. We plan to examine the regional manifestations of these circulation trends in future work.

Code and data availability. Code to calculate the PSI500, and USFC metrics is freely available from the TropD package (https://doi.org/10.5281/zenodo.1157043). CMIP5 and CMIP6 model output is freely available from the Lawrence Livermore National Laboratory (https://esgf-node.llnl.gov/search/cmip5/; https://esgf-node.llnl.gov/search/cmip6/). ERA-Interim and ERA5 reanalysis data are freely available from the European Centre for Medium-Range Weather Forecasts (https://apps.ecmwf.int/datasets/data/interim-full-moda/; https://cds.climate.copernicus.eu/#/search?text=ERA5&type=dataset). MERRA-2 reanalysis data are freely available from NASA (https://doi.org/10.5067/AP1B0BA5PD2K; https://doi.org/10.5067/2E096JV59PK7). JRA-55, CFSR, and CFSv2 reanalysis data are freely available from the National Center for Atmospheric Research Computational and Information Systems Laboratory Research Data Archive (https://doi.org/10.5065/D60G3H5B; https://doi.org/10.5065/D6DN438J; https://doi.org/10.5065/D69021ZF).

Author contribution. KG and SD designed the project, KG performed the formal analysis, and KG and SD prepared the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank Penelope Maher, Karen Rosenlof, and two anonymous reviewers for helpful comments. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

References

- Adam, O., Grise, K. M., Staten, P., Simpson, I. R., Davis, S. M., Davis, N. A., Waugh, D. W., Birner, T., and Ming, A.: The TropD software package (v1): Standardized methods for calculating tropical-width diagnostics, *Geosci. Model Dev.*, 11, 4339–4357, https://doi.org/10.5194/gmd-11-4339-2018, 2018.
- Allen, R. J., and Ajoku, O.: Future aerosol reductions and widening of the northern tropical belt, *J. Geophys. Res.-Atmos.*, 121, 6765–6786, https://doi.org/10.1002/2016JD024803, 2016.

Deleted: , USFC,

Deleted: eddy-driven jet

- Allen, R. J., and Kovilakam, M.: The role of natural climate variability in recent tropical expansion, *J. Climate*, 30, 6329-6350, <https://doi.org/10.1175/JCLI-D-16-0735.1>, 2017.
- 555 Allen, R. J., Sherwood, S. C., Norris, J. R., and Zender, C. S.: Recent Northern Hemisphere tropical expansion primarily driven by black carbon and tropospheric ozone, *Nature*, 485, 350-354, <https://doi.org/10.1038/nature11097>, 2012.
- Allen, R. J., Norris, J. R., and Kovilakam, M.: Influence of anthropogenic aerosols and the Pacific Decadal Oscillation on tropical belt width, *Nat. Geosci.*, 7, 270-274, <https://doi.org/10.1038/ngeo2091>, 2014.
- Amaya, D. J., Siler, N., Xie, S., and Miller, A. J.: The interplay of internal and forced modes of Hadley Cell expansion: lessons from the global warming hiatus, *Clim. Dynam.*, 51, 305-319, <https://doi.org/10.1007/s00382-017-3921-5>, 2018.
- 560 Birner, T., Davis, S. M., and Seidel, D. J.: The changing width of Earth's tropical belt, *Phys. Today*, 67, 38-44, <https://doi.org/10.1063/PT.3.2620>, 2014.
- Choi, J., Son, S.-W., Lu, J., and Min, S.-K.: Further observational evidence of Hadley cell widening in the Southern Hemisphere, *Geophys. Res. Lett.*, 41, 2590-2597, <https://doi.org/10.1002/2014GL059426>, 2014.
- 565 [Choi, J., Son, S.-W., and Park, R. J.: Aerosol versus greenhouse gas impacts on Southern Hemisphere general circulation changes, *Clim. Dynam.*, 52, 4127-4142, <https://doi.org/10.1007/s00382-018-4370-5>, 2019.](https://doi.org/10.1002/2014GL059426)
- Cook, K. H., and Vizy, E. K.: Expanding width of the tropics: Impacts on the ocean, *US CLIVAR Variations*, 16(2), 27-32, <https://doi.org/10.5065/D69Z93QF>, 2018.
- Davis, N., and Birner, T.: On the discrepancies in tropical belt expansion between reanalyses and climate models and among tropical belt width metrics, *J. Climate*, 30, 1211-1231, <https://doi.org/10.1175/JCLI-D-16-0371.1>, 2017.
- 570 Davis, N. A., and Davis, S. M.: Reconciling Hadley cell expansion trend estimates in reanalyses, *Geophys. Res. Lett.*, 45, 11439-11446, <https://doi.org/10.1029/2018GL079593>, 2018.
- Davis, N. A., Seidel, D. J., Birner, T., Davis, S. M., and Tilmes, S.: Changes in the width of the tropical belt due to simple radiative forcing changes in the GeoMIP simulations, *Atmos. Chem. Phys.*, 16, 10083-10095, <https://doi.org/10.5194/acp-16-10083-2016>, 2016.
- 575 Davis, S. M., and Rosenlof, K. H.: A multidagnostic intercomparison of tropical-width time series using reanalyses and satellite observations, *J. Climate*, 25, 1061-1078, <https://doi.org/10.1175/JCLI-D-11-00127.1>, 2012.
- Davis, S. M., Hassler, B., and Rosenlof, K. H.: Revisiting ozone measurements as an indicator of tropical width, *Progress in Earth and Planetary Science*, 5, 56, <https://doi.org/10.1186/s40645-018-0214-5>, 2018.
- 580 Dee, D. P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553-597, <https://doi.org/10.1002/qj.828>, 2011.
- 585

- Dima, I. M., and Wallace, J. M.: On the seasonality of the Hadley Cell, *J. Atmos. Sci.*, 60, 1522–1527, [https://doi.org/10.1175/1520-0469\(2003\)060<1522:OTSOTH>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<1522:OTSOTH>2.0.CO;2), 2003.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- 590 Feng, S., and Fu, Q.: Expansion of global drylands under a warming climate, *Atmos. Chem. Phys.*, 13, 10081–10094, <https://doi.org/10.5194/acp-13-10081-2013>, 2013.
- Forster, P. M., Maycock, A. C., McKenna, C. M., and Smith, C. J.: Latest climate models confirm need for urgent mitigation, *Nat. Clim. Change*, <https://doi.org/10.1038/s41558-019-0660-0>, 2019.
- 595 Frierson, D. M. W., Lu, J., and Chen, G.: Width of the Hadley cell in simple and comprehensive general circulation models, *Geophys. Res. Lett.*, 34, L18804, <https://doi.org/10.1029/2007GL031115>, 2007.
- Fu, Q., Johanson, C. M., Wallace, J. M., and Reichler, T.: Enhanced mid-latitude tropospheric warming in satellite measurements, *Science*, 312, 1179–1179, <https://doi.org/10.1126/science.1125566>, 2006.
- Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegmeier, S., Chabrilat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, *Atmos. Chem. Phys.*, 17, 1417–1452, <https://doi.org/10.5194/acp-17-1417-2017>, 2017.
- 600 Garfinkel, C. I., Waugh, D. W., and Polvani, L. M.: Recent Hadley cell expansion: The role of internal atmospheric variability in reconciling modeled and observed trends, *Geophys. Res. Lett.*, 42, 10824–10831, <https://doi.org/10.1002/2015GL066942>, 2015.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *J. Climate*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>, 2017.
- 610 Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Matthes, K., Santer, B. D., Stone, D., and Tebaldi, C.: The Detection and Attribution Model Intercomparison Project (DAMIP v.1.0) contribution to CMIP6, *Geosci. Model Dev.*, 9, 3685–3697, <https://doi.org/10.5194/gmd-9-3685-2016>, 2016.
- 615 Grise, K. M., and Polvani, L. M.: The response of mid-latitude jets to increased CO₂: Distinguishing the roles of sea surface temperature and direct radiative forcing, *Geophys. Res. Lett.*, 41, 6863–6871, <https://doi.org/10.1002/2014GL061638>, 2014.

- 620 Grise, K. M., and Polvani, L. M.: Is climate sensitivity related to dynamical sensitivity?, *J. Geophys. Res.-Atmos.*, 121, 5159–5176, <https://doi.org/10.1002/2015JD024687>, 2016.
- Grise, K. M., Davis, S. M., Staten, P. W., and Adam, O.: Regional and seasonal characteristics of the recent expansion of the tropics, *J. Climate*, 31, 6839–6856, <https://doi.org/10.1175/JCLI-D-18-0060.1>, 2018.
- Grise, K. M., Davis, S. M., Simpson, I. R., Waugh, D. W., Fu, Q., Allen, R. J., Rosenlof, K. H., Ummenhofer, C. C., Karnauskas, K. B., Maycock, A. C., Quan, X.-W., Birner, T., and Staten, P. W.: Recent tropical expansion: Natural variability or forced response?, *J. Climate*, 32, 1551–1571, <https://doi.org/10.1175/JCLI-D-18-0444.1>, 2019.
- 625 Hawkins, E., and Sutton, R.: Time of emergence of climate signals, *Geophys. Res. Lett.*, 39, L01702, <https://doi.org/10.1029/2011GL050087>, 2012.
- Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, *ECMWF Newsletter*, 159, 17–24, <https://doi.org/10.21957/vf291hehd7>, 2019.
- 630 Hoskins, B. J., Yang, G.-Y., and Fonseca, R. M.: The detailed dynamics of the June–August Hadley Cell, *Q. J. Roy. Meteor. Soc.*, in press, <https://doi.org/10.1002/qj.3702>, 2019.
- Hu, Y., and Fu, Q.: Observed poleward expansion of the Hadley circulation since 1979, *Atmos. Chem. Phys.*, 7, 5229–5236, <https://doi.org/10.5194/acp-7-5229-2007>, 2007.
- 635 Hu, Y., Tao, L., and Liu, J.: Poleward expansion of the Hadley circulation in CMIP5 simulations, *Adv. Atmos. Sci.*, 30, 790–795, <https://doi.org/10.1007/s00376-012-2187-4>, 2013.
- [Hudson, R. D., Andrade, M. F., Follette, M. B., and Frolov, A. D.: The total ozone field separated into meteorological regimes – Part II: Northern Hemisphere mid-latitude total ozone trends, *Atmos. Chem. Phys.*, 6, 5183–5191, <https://doi.org/10.5194/acp-6-5183-2006>, 2006.](https://doi.org/10.5194/acp-6-5183-2006)
- 640 Johanson, C. M., and Fu, Q.: Hadley cell widening: Model simulations versus observations, *J. Climate*, 22, 2713–2725, <https://doi.org/10.1175/2008JCLI2620.1>, 2009.
- Kang, S. M., Polvani, L. M., Fyfe, J. C., and Sigmond, M.: Impact of polar ozone depletion on subtropical precipitation, *Science*, 332, 951–954, <https://doi.org/10.1126/science.1202131>, 2011.
- [Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett.*, 40, 1194–1199, <https://doi.org/10.1002/grl.50256>, 2013.](https://doi.org/10.1002/grl.50256)
- 645 Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., et al.: The JRA-55 Reanalysis: General specifications and basic characteristics, *J. Meteorol. Soc. Jpn. Ser. II*, 93, 5–48, <https://doi.org/10.2151/jmsj.2015-001>, 2015.
- Kossin, J., Emanuel, K., and Vecchi, G.: The poleward migration of the location of tropical cyclone maximum intensity, *Nature*, 509, 349–352, <https://doi.org/10.1038/nature13278>, 2014.
- 650 Kovilakam, M., and Mahajan, S.: Black carbon aerosol-induced Northern Hemisphere tropical expansion, *Geophys. Res. Lett.*, 42, 4964–4972, <https://doi.org/10.1002/2015GL064559>, 2015.

Levine, X. J., and Schneider, T.: Response of the Hadley Circulation to climate change in an aquaplanet GCM coupled to a simple representation of ocean heat transport, *J. Atmos. Sci.*, 68, 769–783, <https://doi.org/10.1175/2010JAS3553.1>, 2011.

Lu, J., Vecchi, G. A., and Reichler, T.: Expansion of the Hadley cell under global warming, *Geophys. Res. Lett.*, 34, L06805, <https://doi.org/10.1029/2006GL028443>, 2007.

655 [Lucas, C., Nguyen, H., and Timbal, B.: An observational analysis of Southern Hemisphere tropical expansion, *J. Geophys. Res.*, 117, D17112, <https://doi.org/10.1029/2011JD017033>, 2012.](#)

[Lucas, C., Timbal, B., and Nguyen, H.: The expanding tropics: a critical assessment of the observational and modeling studies, *WIREs Clim Change*, 5, 89–112, <https://doi.org/10.1002/wcc.251>, 2014.](#)

660 [Maher, P., Kelleher, M. E., Sansom, P. G., and Methven, J.: Is the subtropical jet shifting poleward?, *Climate Dyn.*, 54, 1741–1759, <https://doi.org/10.1007/s00382-019-05084-6>, 2020.](#)

Mantsis, D. F., Sherwood, S., Allen, R., and Shi, L.: Natural variations of tropical width and recent trends, *Geophys. Res. Lett.*, 44, 3825–3832, <https://doi.org/10.1002/2016GL072097>, 2017.

McLandress, C., Shepherd, T. G., Scinocca, J. F., Plummer, D. A., Sigmond, M., Jonsson, A. I., and Reader, M. C.: Separating the dynamical effects of climate change and ozone depletion. Part II: Southern Hemisphere troposphere, *J. Climate*, 24, 1850–1868, <https://doi.org/10.1175/2010JCLI3958.1>, 2011.

665 Min, S.-K., and Son, S.-W.: Multimodel attribution of the Southern Hemisphere Hadley cell widening: Major role of ozone Depletion, *J. Geophys. Res.-Atmos.*, 118, 3007–3015, <https://doi.org/10.1002/jgrd.50232>, 2013.

[Nguyen, H., Lucas, C., Evans, A., Timbal, B., and Hanson, L.: Expansion of the Southern Hemisphere Hadley cell in response to greenhouse gas forcing, *J. Climate*, 28, 8067–8077, <https://doi.org/10.1175/JCLI-D-15-0139.1>, 2015.](#)

670 Polvani, L. M., Waugh, D. W., Correa, G. J. P., and Son, S.: Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere, *J. Climate*, 24, 795–812, <https://doi.org/10.1175/2010JCLI3772.1>, 2011.

Quan, X.-W., Hoerling, M. P., Perlwitz, J., and Diaz, H. F.: On the time of emergence of tropical width change, *J. Climate*, 675 31, 7225–7236, <https://doi.org/10.1175/JCLI-D-18-0068.1>, 2018.

Ryakaczewski, R. R., Dunne, J. P., Sydeman, W. J., García-Reyes, M., Black, B. A., and Bograd, S. J.: Poleward displacement of coastal upwelling-favorable winds in the ocean's eastern boundary currents through the 21st century, *Geophys. Res. Lett.*, 42, 6424–6431, <https://doi.org/10.1002/2015GL064694>, 2015.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., 680 Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y., Chuang, H., Juang, H. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP Climate Forecast System Reanalysis, *B. Am. Meteorol. Soc.*, 91, 1015–1058, 685 <https://doi.org/10.1175/2010BAMS3001.1>, 2010.

- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: The NCEP Climate Forecast System Version 2, *J. Climate*, 27, 2185-2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>, 2014.
- 690 Scheff, J., and Frierson, D. M. W.: Robust future precipitation declines in CMIP5 largely reflect the poleward expansion of model subtropical dry zones, *Geophys. Res. Lett.*, 39, L18704, <https://doi.org/10.1029/2012GL052910>, 2012.
- Schmidt, D. F., and Grise, K. M.: The response of local precipitation and sea level pressure to Hadley Cell expansion, *Geophys. Res. Lett.*, 44, 10573-10582, <https://doi.org/10.1002/2017GL075380>, 2017.
- [Schwendike, J., Govekar, P., Reeder, M. J., Wardle, R., Berry, G. J., and Jakob, C.: Local partitioning of the overturning circulation in the tropics and the connection to the Hadley and Walker circulations, *J. Geophys. Res. Atmos.*, 119, 1322-1339, <https://doi.org/10.1002/2013JD020742>, 2014.](https://doi.org/10.1002/2013JD020742)
- 695 Seidel, D. J., and Randel, W. J.: Recent widening of the tropical belt: Evidence from tropopause observations, *J. Geophys. Res.-Atmos.*, 112, D20113, <https://doi.org/10.1029/2007JD008861>, 2007.
- Seidel, D. J., Fu, Q., Randel, W. J., and Reichler, T. J.: Widening of the tropical belt in a changing climate, *Nat. Geosci.*, 1, 21-24, <https://doi.org/10.1038/ngeo.2007.38>, 2008.
- 700 Sharmila, S., and Walsh, K. J. E.: Recent poleward shift of tropical cyclone formation linked to Hadley cell expansion, *Nat. Clim. Change*, 8, 730-736, <https://doi.org/10.1038/s41558-018-0227-5>, 2018.
- Shaw, T., and Voigt, A.: Tug of war on summertime circulation between radiative forcing and sea surface warming, *Nat. Geosci.*, 8, 560-566, <https://doi.org/10.1038/ngeo2449>, 2015.
- Solomon, A., and Polvani, L. M.: Highly significant responses to anthropogenic forcings of the midlatitude jet in the Southern Hemisphere, *J. Climate*, 29, 3463-3460, <https://doi.org/10.1175/JCLI-D-16-0034.1>, 2016.
- 705 Solomon, A., Polvani, L. M., Waugh, D. W., and Davis, S. M.: Contrasting upper and lower atmospheric metrics of tropical expansion in the Southern Hemisphere, *Geophys. Res. Lett.*, 43, 10496-10503, <https://doi.org/10.1002/2016GL070917>, 2016
- Son, S.-W., Gerber, E. P., Perlwitz, J., Polvani, L. M., Gillett, N. P., Seo, K.-H., Eyring, V., Shepherd, T. G., Waugh, D., Akiyoshi, H., Austin, J., Baumgaertner, A., Bekki, S., Braesicke, P., Brühl, C., Butchart, N., Chipperfield, M. P., Cugnet, D., Dameris, M., Dhomse, S., Frith, S., Garny, H., Garcia, R., Hardiman, S. C., Jöckel, P., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Nakamura, T., Morgenstern, O., Pitari, G., Plummer, D. A., Pyle, J., Rozanov, E., Scinocca, J. F., Shibata, K., Smale, D., Teysseèdre, H., Tian, W., and Yamashita, Y.: Impact of stratospheric ozone on Southern Hemisphere circulation change: A multimodel assessment, *J. Geophys. Res.-Atmos.*, 115, D00M07, <https://doi.org/10.1029/2010JD014271>, 2010.
- 715 Staten, P. W., Lu, J., Grise, K. M., Davis, S. M., and Birner, T.: Re-examining tropical expansion, *Nat. Clim. Change*, 8, 768-775, <https://doi.org/10.1038/s41558-018-0246-2>, 2018.

720

[Staten, P. W., Grise, K. M., Davis, S. M., Karnauskas, K., and Davis, N.: Regional widening of tropical overturning: Forced change, natural variability, and recent trends, *J. Geophys. Res.-Atmos.*, **124**, 6104–6119, <https://doi.org/10.1029/2018JD030100>, 2019.](https://doi.org/10.1029/2018JD030100)

Studholme, J., and Gulev, S.: Concurrent changes to Hadley circulation and the meridional distribution of tropical cyclones, *J. Climate*, **31**, 4367–4389, <https://doi.org/10.1175/JCLI-D-17-0852.1>, 2018.

Tandon, N., Gerber, E., Sobel, A., and Polvani, L.: Understanding Hadley Cell expansion versus contraction: Insights from simplified models and implications for recent observations, *J. Climate*, **26**, 4304–4321, <https://doi.org/10.1175/JCLI-D-12-00598.1>, 2013.

725

Tao, L., Hu, Y., and Liu, J.: Anthropogenic forcing on the Hadley circulation in CMIP5 simulations, *Clim. Dynam.*, **46**, 3337–3350, <https://doi.org/10.1007/s00382-015-2772-1>, 2016.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.

730

Thomas, J. L., Waugh, D. W., and Gnanadesikan, A.: Southern Hemisphere extratropical circulation: Recent trends and natural variability, *Geophys. Res. Lett.*, **42**, 5508–5515, <https://doi.org/10.1002/2015GL064521>, 2015.

Watt-Meyer, O., Frierson, D. M. W., and Fu, Q.: Hemispheric asymmetry of tropical expansion under CO2 forcing, *Geophys. Res. Lett.*, **46**, 9231–9240, <https://doi.org/10.1029/2019GL083695>, 2019.

Waugh, D. W., Garfinkel, C. I., and Polvani, L. M.: Drivers of the recent tropical expansion in the Southern Hemisphere:

735

Changing SSTs or ozone depletion?, *J. Climate*, **28**, 6581–6586, <https://doi.org/10.1175/JCLI-D-15-0138.1>, 2015.

Waugh, D. W., Grise, K. M., Seviour, W. J. M., Davis, S. M., Davis, N., Adam, O., Son, S.-W., Simpson, I. R., Staten, P. W., Maycock, A. C., Ummenhofer, C. C., Birner, T., and Ming, A.: Revisiting the relationship among metrics of tropical expansion, *J. Climate*, **31**, 7565–7581, <https://doi.org/10.1175/JCLI-D-18-0108.1>, 2018.

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Siebesma, A. P., Skinner, C. B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, *Geosci. Model Dev.*, **10**, 359–384, <https://doi.org/10.5194/gmd-10-359-2017>, 2017.

740

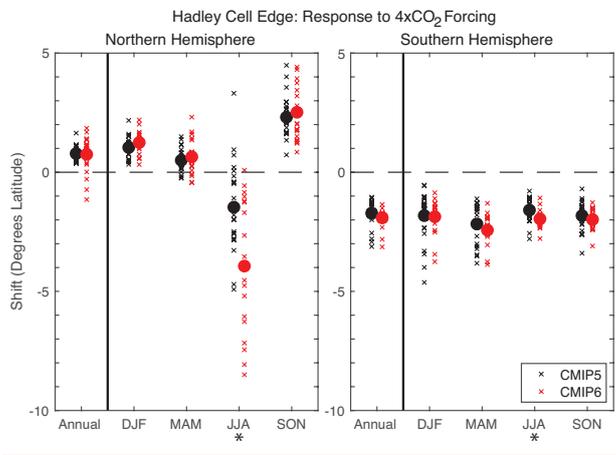
[Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of higher climate sensitivity in CMIP6 models, *Geophys. Res. Lett.*, **47**, e2019GL085782, <https://doi.org/10.1029/2019GL085782>, 2020.](https://doi.org/10.1029/2019GL085782)

745

[Zhou, W., Xie, S.-P., and Yang, D.: Enhanced equatorial warming causes deep-tropical contraction and subtropical monsoon shift, *Nat. Clim. Change*, **9**, 834–839, <https://doi.org/10.1038/s41558-019-0603-9>, 2019.](https://doi.org/10.1038/s41558-019-0603-9)

750

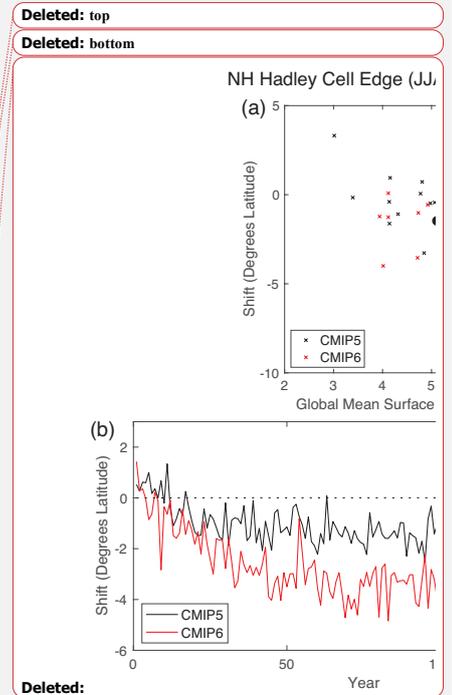
755

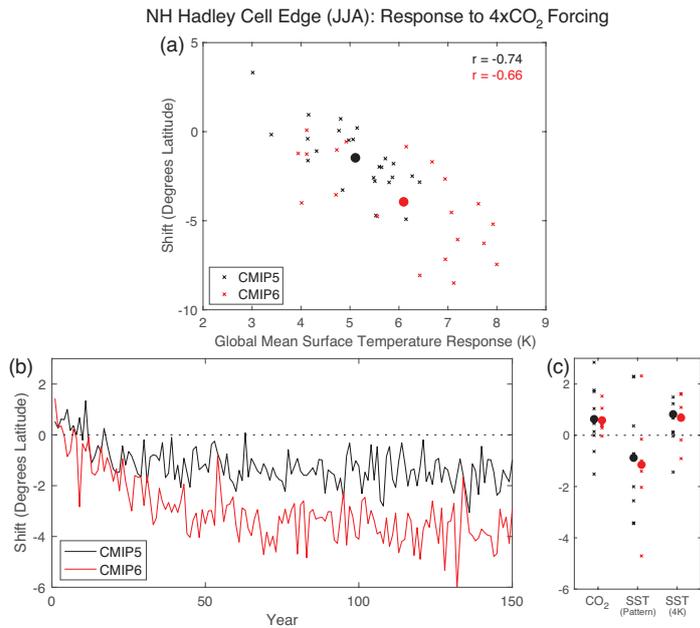


760 Figure 1: Response of (left) NH and (right) SH Hadley cell edge latitude (as measured by PSI500 metric) to 4xCO₂ forcing for (black) CMIP5 and (red) CMIP6 models. Here, the response is defined as the difference in the Hadley cell edge latitude between its mean position during the last 50 years (years 101-150) of the abrupt 4xCO₂ run and its mean position in the pre-industrial control run. The response of each model is shown with a small “x”, and the multi-model mean response is shown as a large dot. Asterisks denote where multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence level via Student’s t-test.

765

770





780 Figure 2: (a) Scatter plot of NH JJA Hadley cell edge response to 4xCO₂ forcing (as measured by PSI500 metric) versus annual-
 785 mean global-mean surface temperature response for (black) CMIP5 and (red) CMIP6 models. (b) Time series of NH JJA Hadley
 cell edge to (first column) quadrupled atmospheric CO₂ concentrations with fixed sea surface temperatures (amip4xCO₂ – AMIP),
 (second column) patterned sea surface temperature increase with fixed atmospheric CO₂ concentrations (amipFuture – AMIP), and
 (third column) uniform 4K sea surface temperature increase with fixed atmospheric CO₂ concentrations (amip4K – AMIP).

Deleted: CO2

Deleted:) and (SST

Deleted: increases

800

805

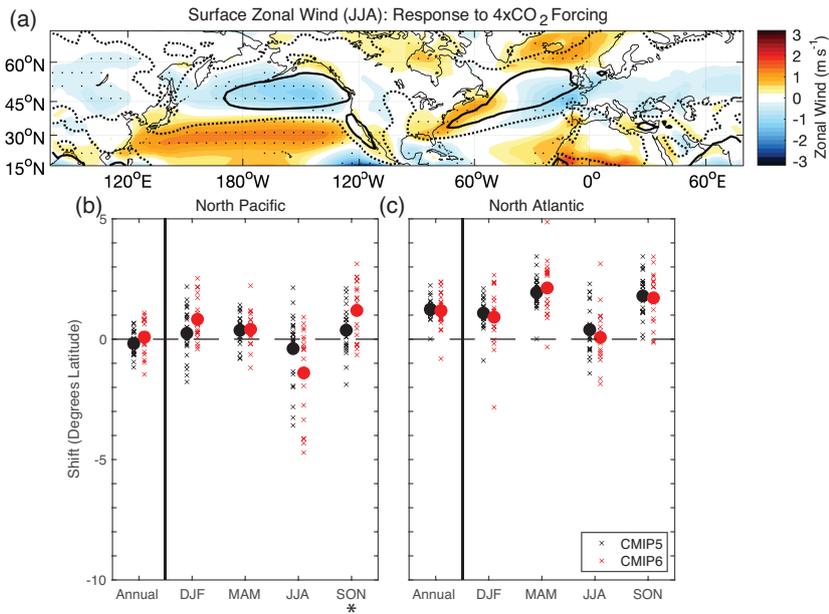
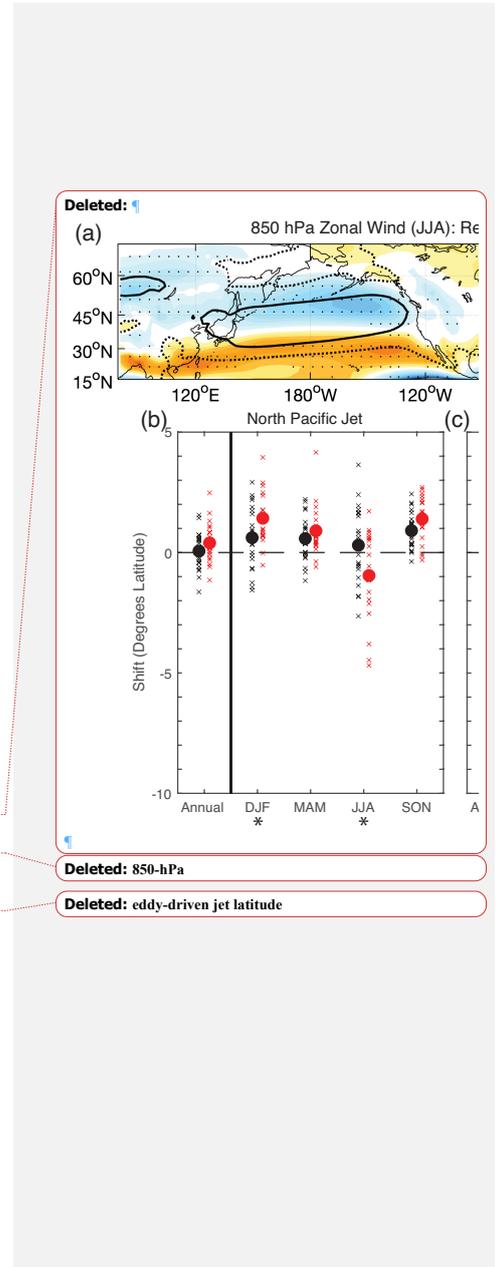


Figure 3: (a) Multi-model mean response of JJA surface zonal wind field to 4xCO₂ forcing for CMIP6 models. The thick dotted and solid lines indicate the 0 m s⁻¹ and 2.5 m s⁻¹ wind contours from the pre-industrial control climatology, respectively. Stippling indicates where the response is statistically significant at the 95% confidence level via Student's t-test. (b, c) As in Fig. 1, but for the USFC metric calculated over the North Pacific (135°E–125°W) and North Atlantic (60°W–0°E) sectors, respectively.

810

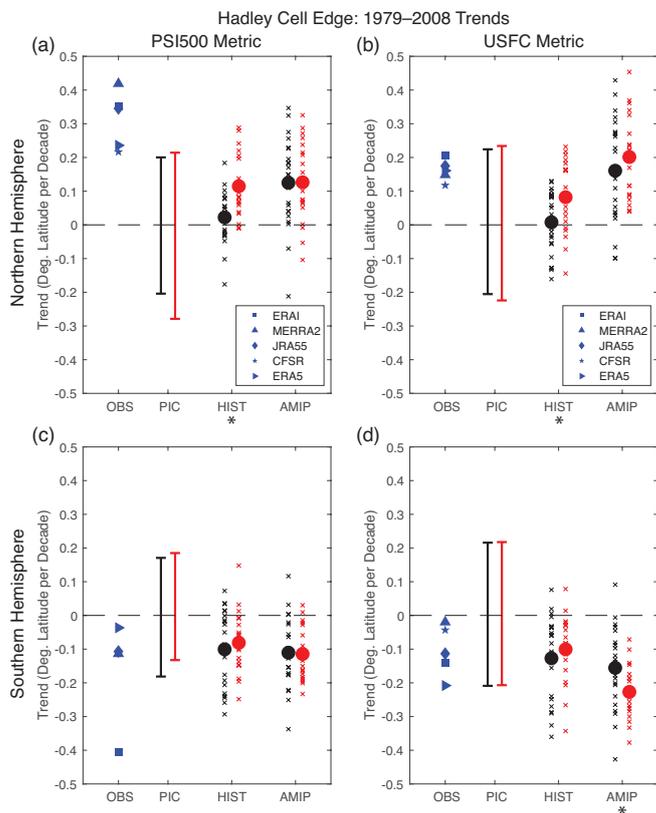
815

820



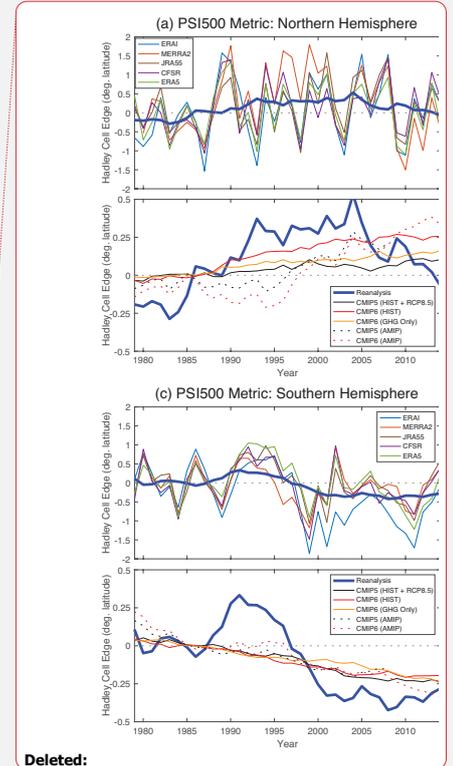
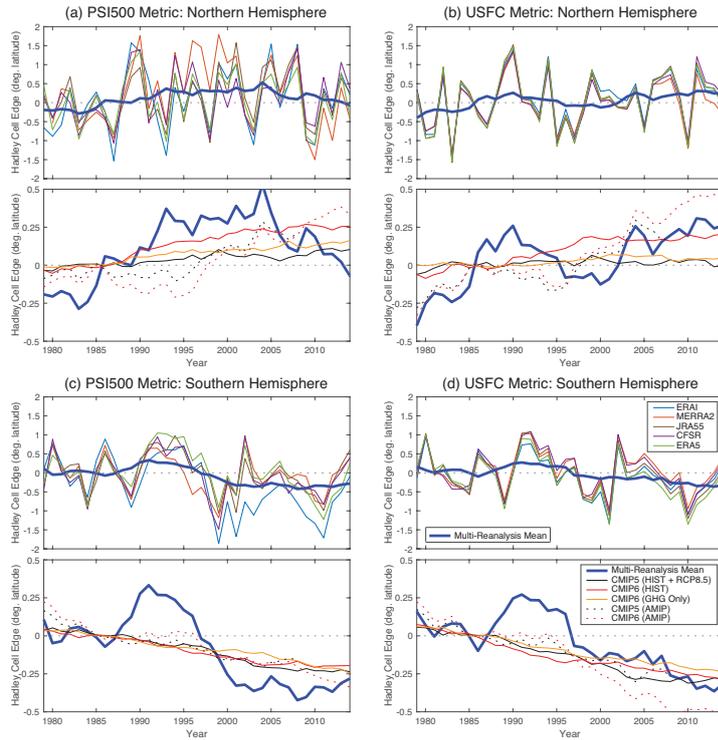
Deleted: 850-hPa

Deleted: eddy-driven jet latitude



835 **Figure 4:** 1979–2008 trends in annual-mean Hadley cell edge latitude, as measured by the (left column) PSI500 and (right column) USFC metrics. Reanalysis trends (OBS, blue symbols) are taken from the ERA-Interim (ERA1), MERRA-2, JRA-55, CFSR, and ERA5 reanalyses. Because MERRA-2 begins in 1980, trends for MERRA-2 are shown for 1980–2008. Control trends (PIC) show the 2.5th–97.5th percentile of trends over 30-year periods from the pre-industrial control runs of (black) CMIP5 and (red) CMIP6 models. Historical trends (HIST) and AMIP trends are calculated from the first ensemble members of (black) CMIP5 and (red) CMIP6 models, where the response of each model is shown with a small “x” and the multi-model mean response is shown as a large dot. Because the historical runs of CMIP5 models end in 2005, they are extended with three years of the RCP 8.5 run until 2008. Asterisks denote where multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence level via Student’s t-test.

840



Deleted:

Deleted: S4

845
 850
 855
 Figure 5: 1979–2014 time series of annual-mean Hadley cell edge latitude, as measured by the (left column) PSI500 and (right column) USFC metrics. All time series are plotted with respect to their 1980–1990 average. The top set of timeseries in each panel shows five reanalyses (ERA-Interim, MERRA-2, JRA-55, CFSR, and ERA5), along with the multi-reanalysis mean (thick blue line). The bottom set of timeseries in each panel shows the multi-reanalysis mean (thick blue line, reproduced from the plot above), as well as the multi-model mean from (black) CMIP5 historical runs (extended with RCP 8.5 until 2014), (red) CMIP6 historical runs, (orange) CMIP6 historical greenhouse gas-only runs (using all available ensemble members; see Table S2), (black dashed) CMIP5 AMIP runs, and (red dashed) CMIP6 AMIP runs. The reanalysis-mean and multi-model mean time series are smoothed with a ten-year running mean to better visualize the low frequency variability in each time series. Note that the scale is different for the top and bottom sets of timeseries in each panel.

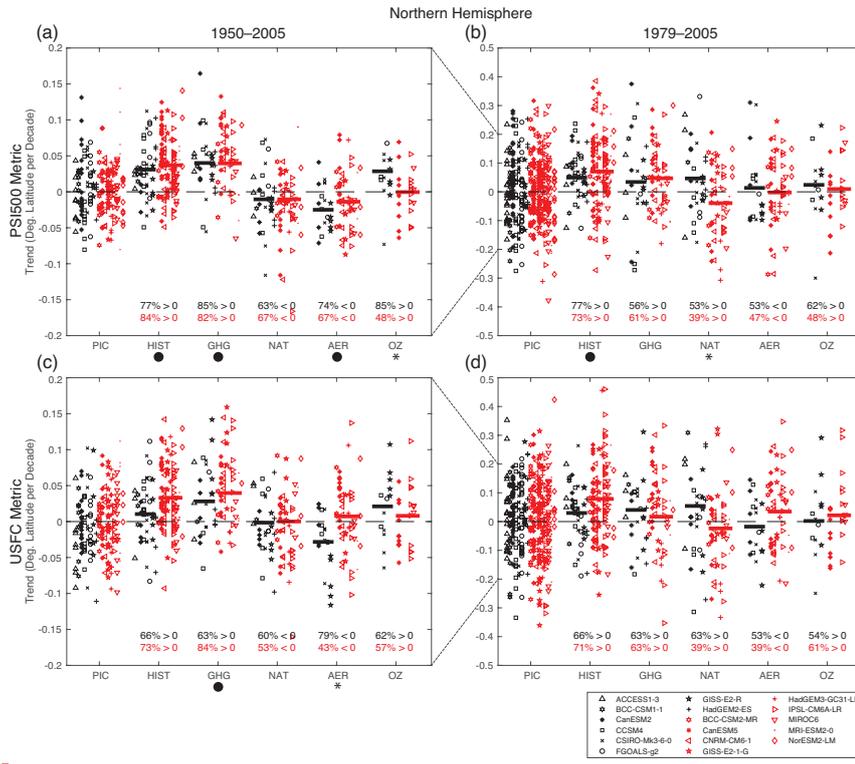
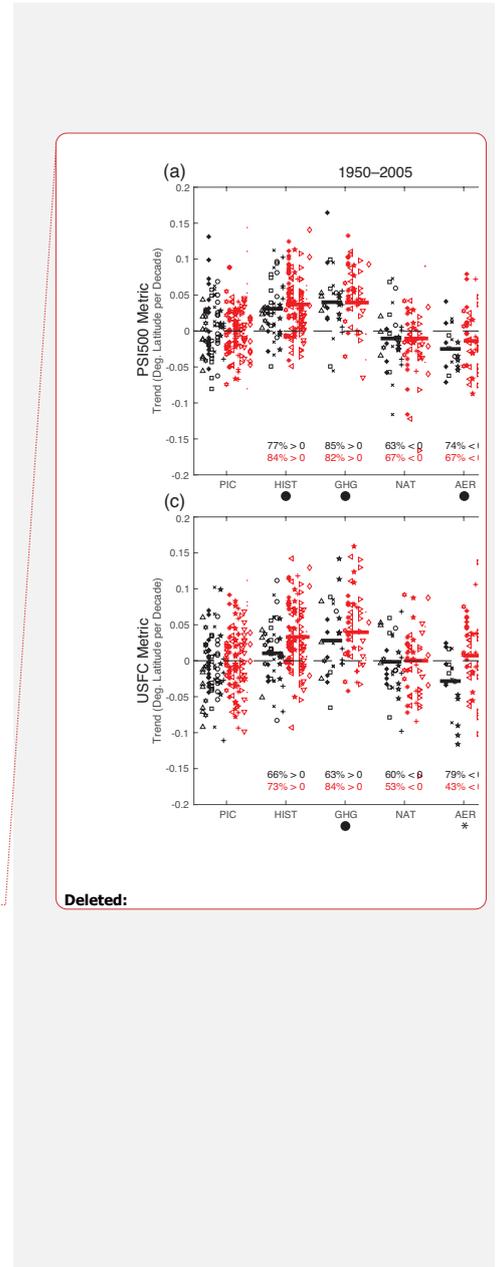
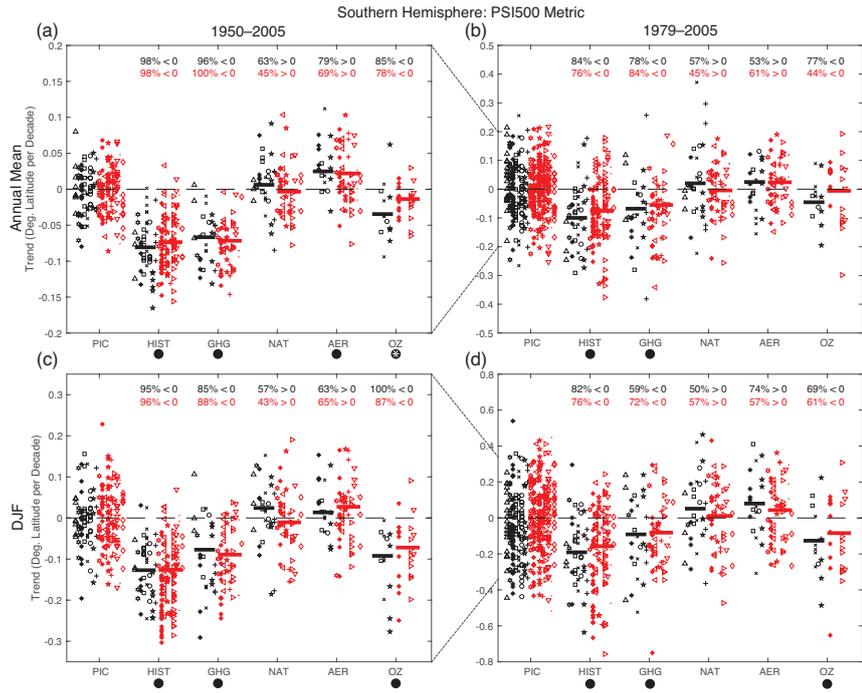


Figure 6: Trends in annual-mean NH Hadley cell edge latitude over (left column) 1950–2005 and (right column) 1979–2005 for (black) CMIP5 and (red) CMIP6 models. Trends are shown for all available ensemble members of the following runs: (HIST) historical, (GHG) greenhouse gas only, (NAT) solar and volcanic only, (AER) anthropogenic aerosol only, and (OZ) ozone only. Multi-model mean trends are shown as thick horizontal lines. Trends are also shown for all independent time periods of equivalent length from the pre-industrial control (PIC) runs. Large black dots denote forcings with trends statistically different from zero in both CMIP5 and CMIP6 models. Asterisks denote where multi-model means of CMIP5 and CMIP6 models are statistically different at the 95% confidence level via Student's *t*-test.



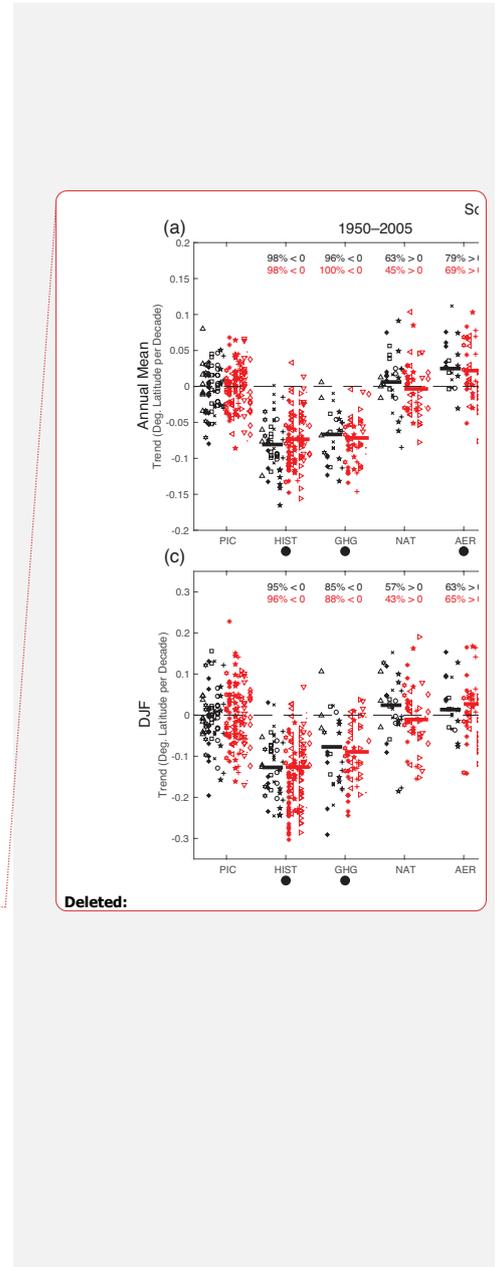


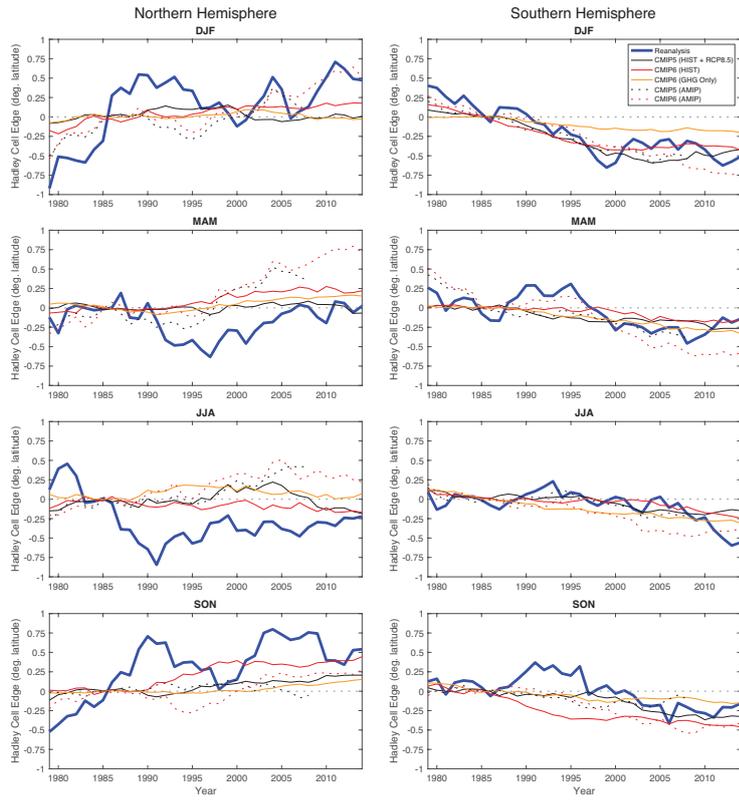
885

Figure 7: As in Fig. 6, but for (top row) annual-mean SH Hadley cell edge latitude and (bottom row) DJF-mean SH Hadley cell edge latitude (as measured by PSI500 metric).

890

895





905 **Figure 8:** As in Fig. 5, but for time series of 1979–2014 seasonal-mean Hadley cell edge latitudes (as measured by USFC metric) in both hemispheres.

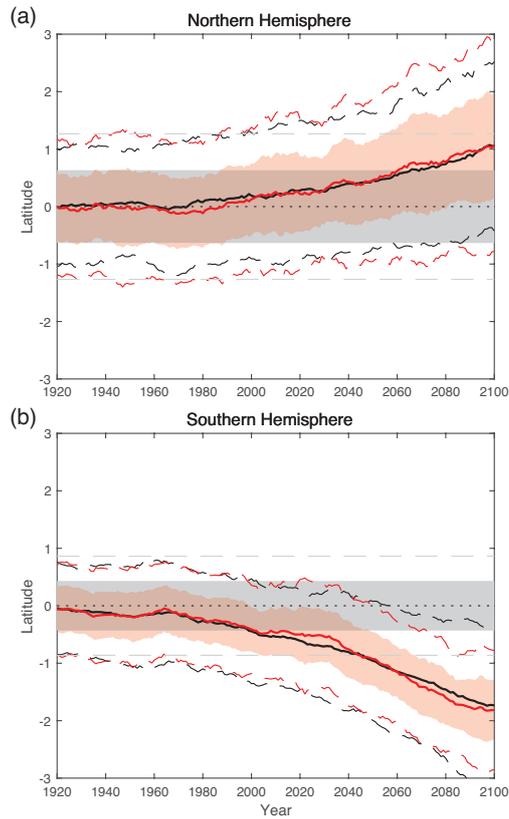


Figure 9: 1920–2100 time series of annual-mean Hadley cell edge latitude (as measured by PSI500 metric) from CMIP5 historical + RCP 8.5 and CMIP6 historical + SSP 5-8.5 runs. The multi-model mean (smoothed with 10-year running mean) for CMIP5 (CMIP6) models is plotted as a solid black (red) line. One standard deviation range across the inter-model spread of CMIP6 models is shown as red shading. Two standard deviation range across the inter-model spread of CMIP5 (CMIP6) models is plotted as black (red) dashed lines. One and two standard deviation ranges about the pre-industrial control latitude of CMIP6 models are shown as gray shading and gray dashed lines, respectively. All latitudes are plotted with respect to the multi-model mean pre-industrial control latitude.

<u>CMIP5 Model</u>	<u>Resolution ($^{\circ}\text{lon} \times ^{\circ}\text{lat}$)</u>	<u>CMIP6 Model</u>	<u>Resolution ($^{\circ}\text{lon} \times ^{\circ}\text{lat}$)</u>
<u>ACCESS1.0</u>	<u>$1.875^{\circ} \times 1.25^{\circ}$</u>	<u>BCC-CSM2-MR[#]</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>
<u>ACCESS1.3</u>	<u>$1.875^{\circ} \times 1.25^{\circ}$</u>	<u>BCC-ESM1</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>
<u>BCC-CSM1.1</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>	<u>CAMS-CSM1-0[#]</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>
<u>BCC-CSM1.1(m)</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>	<u>CanESM5[#]</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>
<u>BNU-ESM</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>	<u>CESM2[#]</u>	<u>$1.25^{\circ} \times 0.9424^{\circ}$</u>
<u>CanESM2 (CanAM4)</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>	<u>CESM2-WACCM[#]</u>	<u>$1.25^{\circ} \times 0.9424^{\circ}$</u>
<u>CCSM4</u>	<u>$1.25^{\circ} \times 0.9424^{\circ}$</u>	<u>CNRM-CM6-1[#]</u>	<u>$1.40625^{\circ} \times 1.4008^{\circ}$</u>
<u>CNRM-CM5</u>	<u>$1.40625^{\circ} \times 1.4008^{\circ}$</u>	<u>CNRM-ESM2-1[#]</u>	<u>$1.40625^{\circ} \times 1.4008^{\circ}$</u>
<u>CSIRO Mk3.6.0</u>	<u>$1.875^{\circ} \times 1.8653^{\circ}$</u>	<u>E3SM-1-0</u>	<u>$1.0^{\circ} \times 1.0^{\circ}$</u>
<u>EC-EARTH</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>	<u>EC-Earth3[#]</u>	<u>$0.7031^{\circ} \times 0.7018^{\circ}$</u>
<u>FGOALS-g2</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>	<u>EC-Earth3-Veg[#]</u>	<u>$0.7031^{\circ} \times 0.7018^{\circ}$</u>
<u>GFDL CM3</u>	<u>$2.5^{\circ} \times 2.0^{\circ}$</u>	<u>GISS-E2-1-G</u>	<u>$2.5^{\circ} \times 2.0^{\circ}$</u>
<u>GISS-E2-R</u>	<u>$2.5^{\circ} \times 2.0^{\circ}$</u>	<u>HadGEM3-GC31-LL</u>	<u>$1.875^{\circ} \times 1.25^{\circ}$</u>
<u>HadGEM2-ES</u> <u>(HadGEM2-A)</u>	<u>$1.875^{\circ} \times 1.25^{\circ}$</u>	<u>IPSL-CM6A-LR[#]</u>	<u>$2.5^{\circ} \times 1.2676^{\circ}$</u>
<u>INM-CM4.0</u>	<u>$2.0^{\circ} \times 1.5^{\circ}$</u>	<u>MIROC6[#]</u>	<u>$1.40625^{\circ} \times 1.4008^{\circ}$</u>
<u>IPSL-CM5A-LR</u>	<u>$3.75^{\circ} \times 1.8947^{\circ}$</u>	<u>MRI-ESM2-0[#]</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>
<u>IPSL-CM5A-MR</u>	<u>$2.5^{\circ} \times 1.2676^{\circ}$</u>	<u>NESM3[#]</u>	<u>$1.875^{\circ} \times 1.8653^{\circ}$</u>
<u>IPSL-CM5B-LR</u>	<u>$3.75^{\circ} \times 1.8947^{\circ}$</u>	<u>NorESM2-LM</u>	<u>$2.5^{\circ} \times 1.8947^{\circ}$</u>
<u>MIROC5</u>	<u>$1.40625^{\circ} \times 1.4008^{\circ}$</u>	<u>SAM0-UNICON</u>	<u>$1.25^{\circ} \times 0.9424^{\circ}$</u>
<u>MIROC-ESM</u>	<u>$2.8125^{\circ} \times 2.7906^{\circ}$</u>	<u>UKESM1-0-LL[#]</u>	<u>$1.875^{\circ} \times 1.25^{\circ}$</u>
<u>MPI-ESM-LR</u>	<u>$1.875^{\circ} \times 1.8653^{\circ}$</u>		
<u>MPI-ESM-MR</u>	<u>$1.875^{\circ} \times 1.8653^{\circ}$</u>		
<u>MRI-CGCM3</u>	<u>$1.125^{\circ} \times 1.1215^{\circ}$</u>		
<u>NorESM1-M</u>	<u>$2.5^{\circ} \times 1.8947^{\circ}$</u>		

930 **Table 1: Global climate models used in this study. Resolution indicates the horizontal resolution at which the data are provided. Bolded models denote those models with output available from the amip4xCO₂, amipFuture/amip-future4K, and amip4K/amip-p4K runs. The # symbols denote the CMIP6 models with output available from the SSP 5-8.5 run. The first ensemble member is used for each model, except in two cases when it is unavailable. In those cases, the ‘r8i1p1f1’ ensemble member is used for the abrupt 4xCO₂ run of EC-Earth3, and the ‘r2i1p1f1’ ensemble member is used for the amip-p4K run of MIROC6.**
935

940

<u>Reanalysis</u>	<u>Resolution ($^{\circ}\text{lon} \times ^{\circ}\text{lat}$)</u>	<u>Time Period</u>	<u>Citation</u>
<u>ERA-5</u>	<u>$0.25^{\circ} \times 0.25^{\circ}$</u>	<u>1979–2014</u>	<u>Hersbach et al. (2019)</u>
<u>ERA-Interim</u>	<u>$0.75^{\circ} \times 0.75^{\circ}$</u>	<u>1979–2014</u>	<u>Dee et al. (2011)</u>
<u>JRA-55</u>	<u>$1.25^{\circ} \times 1.25^{\circ}$</u>	<u>1979–2014</u>	<u>Kobayashi et al. (2015)</u>
<u>NASA MERRA-2</u>	<u>$0.625^{\circ} \times 0.5^{\circ}$</u>	<u>1980–2014</u>	<u>Gelaro et al. (2017)</u>
<u>NCEP CFSR</u>	<u>$0.5^{\circ} \times 0.5^{\circ}$</u>	<u>1979–2010</u>	<u>Saha et al. (2010)</u>
<u>_____ CFSv2</u>		<u>2011–2014</u>	<u>Saha et al. (2014)</u>

Table 2: Reanalysis data sets used in this study.

945

950

955

960

	Northern Hemisphere	Southern Hemisphere
Annual	0.28	0.56
	0.09	0.58
DJF	0.72	0.39
	0.55	0.51
MAM	0.15	0.55
	-0.44	0.64
JJA	-0.74	0.59
	-0.66	0.75
SON	0.40	0.44
	0.61	0.37

Table 3: Correlations between poleward shift of Hadley cell edge latitude in response to $4\times\text{CO}_2$ forcing (as measured by PSI500 metric) with annual-mean global-mean surface temperature response to $4\times\text{CO}_2$ forcing. Positive correlations imply that models that warm more shift the Hadley cell edge further poleward. Correlations for CMIP5 and CMIP6 models are shown in the top and bottom rows of each cell, respectively. Correlations that are statistically significant at the 95% confidence level via Student's t-test are bolded.

Deleted: 1